EVALUATING THE EFFECTIVENESS OF STATE R&D TAX CREDITS

by

Yujeung Ho

Bachelor of Science, Yonsei University, 1994

Master of City Planning, Seoul National University, 1996

Submitted to the Graduate Faculty of

Graduate School of Pulbic and Intertional Affairs in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2006

UNIVERSITY OF PITTSBURGH

Graduate School of Public and International Affairs

This dissertation was presented

by Yujeung Ho

It was defended on

June 21, 2006

and approved by

Dr. Siddharth Chandra, Associate Professor, GSPIA, University of Pittsburgh

Dr. David Miller, Associate Professor, GSPIA, University of Pittsburgh

Dr. Vijai Singh, Vice Chancellor, University of Pittsburgh

Dissertation Advisor: Dr. Sabina Deitrick, Associate Professor, GSPIA, University of Pittsburgh

EVALUATING THE EFFECTIVENESS OF STATE R&D TAX CREDITS

Yujeung Ho, PhD

University of Pittsburgh, 2006

This paper aimed to analyze the effectiveness of state R&D tax credit programs in the context of R&D-relevant policies and regional economic development policies. Although there were extensive theoretical recommendations for promoting private R&D, and state R&D tax credit programs have been one of the most popular regional economic development programs, only few evaluations of state R&D tax credit programs have been conducted. Inspired by this lack of previous study, this study provided an empirical finding for the effectiveness of these programs by applying a quasi-experimental approach, which means conducting experiments without randomness, for comparing states with tax credits and states with no credits.

For dealing with the embedded non-randomness, plausible other explanations that weaken the causal relationship between the programs and the effects were examined and ruled out as much as possible. Rival hypotheses were selected using different tax and government policies, overall business and R&D-specific environments, and firm characteristics. They were eliminated by constructing valid control groups, using the difference-in-differences and matching methods, selecting covariates and matching variables as observable variables, and absorbing year-specific fixed effects and cross-sectional-fixed effects as unobservable variables. The decision was made based on multiple estimates and multiple datasets. The research analyzed two sets of industries: the all industry group and high-technology industy.

The major findings are : 1) state R&D tax credits positively affect the increase in R&D spending and increase in employment; 2) positive effects on R&D spending are widespread across the all industry group while positive effects on employment are limited to high-technology industry overall; 3) positive effects on R&D spending are also spread out to different sized firms in both the all industry group and high-technology industry; and 4) positive effects on employment are found mainly in large firms in both the all industry group and high-technology industry.

The above findings support the utilization of state R&D tax credits. As an indirect intervention, state R&D tax credit programs can increase productivity and encourage innovation by generating additional private R&D activities. State R&D tax credit programs can also make a positive contribution to regional economic growth through the growth of R&D-relevant and high-technology industries.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank my advisory committee for their help and support. Dr. Chandra, Dr. Miller, and Dr. Singh gave me helpful advices and comments from setting up my research question to suggesting policy implications. Especially, I cannot express enough gratitude for the support and encouragement by Dr. Deitrick. She was the one who always on my side and encouraged me to be a social scientist who can contribute to the world. I also would like to say special thanks to Dr. Stephens who taught me econometrics.

I would like to thanks my parents for their endless love and support. My father, Ki-Jun Ho and my mother, Hak-Ja Kim, always encouraged and inspired me to make accomplishments including this work. I'm always in their debt. I also would like to share the joy with my brother, Sung-Jin Ho, and sister, Eugene Ho. I'm blessed with the brotherhood we had and will have throughout our life. I'm also grateful to my husband, Jiwoong Yoon, and the greatest gift I've ever received, Jungwon (Jenny) and Sungjoon (Andy).

Moreover, I also feel lucky for having a chance to build friendship with people I met in Pittsburgh. There are too many friends and families to thank for sharing their heart and time with me and my family.

Finally, this thesis is dedicated to my father, Ki-Jun Ho who gave me all his support and love which are more than I can say. I regret that he cannot share this moment with me, but I believe that he would be happy. I love you, dad.

1.0 INTRODUCTION

This paper analyzes the effectiveness of state R&D tax credit programs in the context of R&D-relevant policies and regional economic development policies. State R&D tax credit programs have been implemented based on extensive theoretical recommendations for promoting private R&D for economic growth. Such programs have also been popular regional economic development initiatives. However, the practical usefulness of these programs is somewhat ambiguous because there has been little research into their effectiveness, especially in quantitative terms. Therefore, this study aims to provide an empirical finding on the effectiveness of state R&D tax credit programs. Empirically testing the effects of state R&D tax credit programs is also meaningful in the sense that the evaluation methods used in this study could be applied to other R&D-relevant policies or other regional economic development policies, which have rarely been evaluated.[1]

Encouraging private R&D activities has been emphasized as a strategy to stimulate economic growth. R&D plays a primary role in increasing productivity and promoting innovation, or technological progress, and consequently accomplishing economic growth in the long run. Long ago, Schumpeter emphasized the role of entrepreneurs and innovators in

---

[1] See Klette, Møen and Griliches (2000) on the lack of evaluation of R&D-relevant policies and Reese and Fasenfest (1999) and Bartik and Bingham (1997) on the lack of evaluation of regional economic development policies.

economic growth by applying the concept of “creative deconstruction,” in which new economic systems are continuously created to replace the old ones (Schumpeter, 1934, 1942).

More recently, endogenous growth theorists explained that the economy grows with the accumulation of R&D or human capital, which are understood as being similar to conventional inputs with some unique aspects (Romer, 1986, 1990; Lucas, 1988). R&D can increase outputs without decreasing returns through productivity growth because R&D can be shared by multiple actors without additional costs, which is recognized as positive externality or spillovers (Arrow, 1962a; Romer, 1990; Spence, 1984). This possible externality also creates valuable benefits for the whole society in the sense that the gains from R&D for society are much higher than those for the private sector. However, this feature also gives the private sector negative incentives to perform R&D to avoid its common utilization without proper compensation. This possibility of under-investing in R&D by the private sector is an important basis for government intervention to support private R&D. It is also believed that spillovers occur with some geographical barriers, which is recognized as localized knowledge spillovers (Lucas, 1993; Von Hipple, 1994; Jaffe, Trajtenberg, and Henderson, 1993). This is another fundamental basis for government support of private R&D, especially at the regional level.

Regional economic development theories also support R&D and R&D-intensive industry as facilitators of regional economic growth. For example, R&D-intensive industry tends to produce high value-added exportable goods.[2] More specifically, such industry tends to develop new products that offer a comparative advantage to a region during the initial development

[2] The importance of exportable goods for regional economic growth is indicated in economic base theory. A detailed discussion follows in Chapter 3, section 3.2.1.

period, before the products are matured and standardized.[3] R&D also plays an important role in producing customized goods or niche market products for satisfying the demand of variety and quality as a newly emerging type of regional economic growth through flexible specialization production systems.[4]

Based on these theoretical recommendations for promoting private R&D, R&D tax credit programs have been widely implemented in the United States. The R&D tax credit was introduced at the federal level by the Economic Recovery Tax Act (ERTA) of 1981. Since then, this tax credit law also has been enacted at the state level, and 38 states adopted their own R&D tax credit programs by 2003.

The R&D tax credit can be described as an indirect government intervention to support private R&D, which can be distinguished from direct government funding or government-performed R&D. The R&D tax credit provides flexible utilization and autonomy to the private sector and therefore does not distort the firm's decision making, subsequently generating efficient resource allocation.

However, questions have been raised about its practical usefulness. These relate to possible windfall effects, undesirable inequality brought about through the exclusion of small firms, and low response elasticity (Klettle, Møen and Griliches, 2000; Sigalla and Viard, 1999). Therefore, with the absence of empirical analyses, and quantitative analysis in particular, the effectiveness of this specific program is still in question. Research to evaluate state R&D tax credit programs has been hindered by some complexities involved in regional level economic

[3] The importance of developing new products within a product development cycle is emphasized in product cycle theory. A detailed discussion follows in Chapter 3, section 3.2.3.

[4] The notion of the necessity of developing specialized and differentiated goods for regional economic growth is found in flexible production theory. A detailed discussion follows in Chapter 3, section 3.2.4.

development programs, including the huge heterogeneity across states in terms of economic environment, tax structure, and tax credit law itself (Hall and Wosinska, 1999). Therefore, only a few studies have evaluated state R&D tax credit programs, compared with large volumes of research on the federal R&D tax credit program. Therefore, it is important to evaluate the effectiveness of state R&D tax credit programs and the practical usefulness of such programs, especially at the regional level.

Social program evaluations generally encounter the difficulty of either finding a counterfactual in reality or randomly manipulating treatments and randomly assigning samples. Accordingly, quasi-experimentation, which means experimentation without randomness, is appropriate for social program evaluations; in this approach for dealing with embedded non-randomness, recognizing other plausible explanations and ruling them out as much as possible is the most important task (Shadish, Cook and Campbell, 2003).

To do this, in other words, to make causal inferences about a program and its effects and making those effects stronger, it is useful to use the theory of validity.[5] In order to increase internal validity, which is involved in establishing causation, it is important to increase similarity and remove differences within observations. However, in order to increase external validity, which is relevant to generalizing the revealed causation, it is important to increase diversity and heterogeneity within observations. Therefore, there is always a trade-off between internal validity and external validity.

[5] The theory of validity was developed by Cook and Campbell (1979) and Shadish, Cook and Campbell (2003). There are four types of validity: internal validity, external validity, construct validity, and statistical conclusion validity. Construct validity is involved in increasing representativeness and statistical validity is involved in statistically increasing causal relationships. A detailed discussion follows in Chapter 5, section 5.1.

Based on the theory of validity, this research uses several evaluation strategies to deal with the heterogeneity across states. These methods are the difference-in-differences method and the matching method for comparing states with tax credits to states without tax credits. The DD method and the matching method estimate the difference in outcome changes between the experimental group and the control group, which indicates a program effect. I analyze the characteristics of state R&D tax credit programs in order to construct valid comparison groups. With recognizing rival hypotheses based on significant determinants for regional economic growth, I choose control variables and matching variables based on these rival hypotheses. I use multiple yearly observations for each comparison group to control for time-specific effects and cross-sectional-fixed-effects as well. Through these evaluation strategies, the systematic differences between comparison groups are reduced, resulting in the increase in internal validity.

Another evaluation strategy applied in this study is providing multiple findings based on multiple estimates and multiple datasets (i.e., state-level data and firm-level data). Multiple estimates are obtained from multiple methods (i.e., the DD method and the matching method), multiple outcome variables (i.e., R&D spending and employment), two industrial level analyses (i.e., all industry and high-technology industry), and level and log analyses. Thus the empirical findings are more reliable and stronger because they are based on multiple estimates, resulting in the increase in external validity.

By using the evaluation strategies mentioned above, this study analyzes the overall effects of state R&D tax credit programs across states and industries rather than limiting the evaluation to specific states or industries. The analysis therefore encompasses the effects on R&D spending and employment and the effects for the all industry and the high-technology industry. Consequently, the findings could provide overall policy implications for state R&D tax credits, even though they do not provide detailed policy recommendations for the specific applications, with a strong external validity.

This study also analyzes the possibility of different effects of state R&D tax credit programs by firm size. R&D activities generally have a large fixed cost. It is therefore, recommended that R&D incentives be specifically targeted to small firms. An interesting and important question for further consideration is whether state R&D tax credit programs can be one of R&D incentives to promote R&D activities for small firms.

Importantly, the applied evaluation methods could be modified and applied to evaluate other regional level economic development programs. As a response of prompt necessity for evaluating regional economic development policies, the advanced evaluation techniques could fill a gap between a theory and its realization.

This paper is organized as follows. Chapter 2 introduces the context and background involved in evaluating regional economic development programs and R&D-relevant programs and also briefly describes the evaluation strategies developed in this study. Chapter 3 explicates economic growth theories focused on the role of technological progress; these theories are valuable for understanding the importance of R&D for economic growth and the necessity of government intervention to support private R&D. This chapter also summarizes regional economic development theories, which are important for understanding the role of R&D in the high-technology industry, especially for regional economic growth. Chapter 4 outlines the historical development of the utilizations of R&D tax credit programs in the United States and analyzes the previous relevant evaluation literatures. Chapter 5 examines possible methodological approaches for evaluating social programs, which requires dealing with the non-equivalent features of treated and non-treated groups and separating program effects from outcome changes that result from other relevant factors. Chapter 6 describes the detailed evaluation strategies and empirical specifications used in this analysis, which are based on a quasi-experimental approach for applying empirical analyses. Chapter 7 empirically tests the effectiveness of R&D tax credits in terms of the increase in R&D spending and employment, and Chapter 8 concludes this study with some policy implications.

2.0 CONTEXT AND BACKGROUND

The program to be investigated in this study is the state R&D tax credit. State R&D tax credit policies are important for two areas: regional level economic development policies and R&D-relevant policies. In order to accomplish regional economic growth, such as employment and output growth, supporting R&D is considered as one of important tasks of regional economic development programs. More broadly, encouraging private R&D activities is one of fundamental goals of public policies in the ground of that it is hard to reach the optimal point of private R&D investments based on market failure while R&D activities can promote economic growth by increasing productivity.[6]

Under this context, researchers mostly encounter some difficulties for evaluating state R&D tax credit programs, more broadly R&D-relevant policy programs. These difficulties include an ambiguous range of effects of regional economic development programs, different features of state R&D tax credit program by state, the involvement of spillover effects, and data availability. To tackle these difficulties and to develop some evaluation strategies are the major task of this study and main points of these strategies are introduced in this chapter. These

---

[6] R&D is believed as a major force for increasing productivity from its externality which means that R&D is sharable by muptiple actors without additional costs as one of public goods (see, for example, Griliches, 1979, 1992; Romer, 1986, 1990). There are some empirical studies supporting this argument, such as Mowery and Rosenberg (1989), Nadiri (1993) and Nelson (2000).

strategies include developing analytical methods and utilizing multiple estimates and datasets for providing the more reliable and strong evidence of the program effect.

2.1 CONTEXT OF STATE R&D TAX CREDITS AND THEIR EVALUATIONS

The R&D tax credit has been popularly implemented in many U.S. states with increasing efforts in economic development to cope with severe inter-regional competition. The R&D tax credit plays a significant role for promoting economic development because (1) supporting R&D is crucial for future economic growth because R&D is considered as the most value-creating activity in terms of technical innovation and job creation[7]; and (2) tax incentives are getting widely utilized from scarce financial resources of local government and relatively diminishing federal support since the 1970s. Within this context, the state R&D tax credit attempts to promote regional economic growth broadly from the primary purpose of encouraging private R&D activity extending to increases in employment and output. Therefore, understanding and evaluating this program can be carried on within these contexts.

In the United States, the R&D tax credit was introduced at the federal level by the Economic Recovery Tax Act (ERTA) of 1981.[8] Since then, this tax credit law also has been implemented at the state level, and 38 states adopted their own R&D tax credit programs by

[7] Rashkin (2003) pointed out that R&D-related jobs are mostly high-wage and professional, and they can create linked opportunities for new manufacturing jobs. Hall and Wosinska (1999) also pointed out that R&D related jobs are high-wage and creates multiplier effects to other workers presumably within the state based on localized spillover effects.

[8] The federal R&D tax credit has been developed and modified over 25 years, and importantly renewed ten times. Its several modifications for better implementation are, for example, redefining eligible R&D activities, changing credit rates or establishing Alternative Incremental Credit (AIC) especially for start-up companies. For the detailed utilization of this law, refer to Hall and Van Reenen (2000) and Rashkin (2003).

2003. Most state-level R&D tax credit programs are modifications of the federal law, although the details differ somewhat.[9] In the international perspective, more than 25 countries have utilized R&D tax credits, as well.[10]

The motivation of this paper comes from the well-recognized lack of evaluation of state R&D tax credit programs, in particular in quantitative terms (see, for example, Hall and Woskinka, 1999; Sigalla and Viard, 1999). Despite extensive evaluations of the federal R&D tax credit policy,[11] there are limited studies on evaluating state R&D tax credit policies. Most of these were relied on descriptive analyses or interview and survey approaches focused on only one or two individual state programs.[12]

Klette, Møen and Griliches (2000) more broadly pointed out the lack of evaluation of R&D-related policies using quantitative methods, in particular, the difference-in-differences method, which is used for five previous studies for evaluating government-sponsored R&D projects,[13] and is used in this dissertation. They also mentioned most available evaluation

---

[9] For example, most of the state R&D tax credits are R&D expense-based as following the federal R&D tax credit, but some state laws are employment-based (for example, Mississippi and Vermont). Indeed, most of them are utilized as an incremental way, while some state laws are utilized as a non-incremental way.

[10] For the detail, see Rashkin (2003).

[11] Office of Technology Assessment (1985) and Hall and Van Reenen (2000) are excellent collections of previous studies on the federal R&D tax credit. According to overall evaluation of empirical studies of assessment of R&E tax credit by Office of Technology Assessment (1985), the estimated sensitivity of R&D spending to the R&D tax rate ranges from 0.3:1 to 0.5:1, that means one dollar tax reduction will increase R&D spending by 30 to 50 cents. From the analysis of Hall and Van Reenen (2000), it is broadly believed approximately one dollar increase in R&D spending results from one dollar increase in the tax credit. Although the difference of estimated sensitivity cannot be easily explained by researchers, there are common conclusions about the statistically significant and positive effect of the federal level R&D tax credit.

[12] Among few, there are a study on California R&D tax credit (Hall and Woskinka, 1999), three consecutive studies on Washington R&D tax credit (the Department of Revenue of Washington State, 1997, 2000, and 2003), and a atudy on Virginia R&D tax credit (Secretaries of Technology and Commerce and Trade of Virginia State, 2000).

[13] These five previous studies are Irwin and Klenow (1996), Lerner (1999), Branstetter and Sakakibara (1998), Griliches and Regev (1998), and Klette and Møen (1999).

studies except the above recent studies used other methods such as case study and interview approaches.

This paucity of evaluation in quantitative ways stems from the following reasons. First, as an economic development strategy, it is hard to answer the question: to what extent the programs affect economic outcomes under the consideration of engaging in multiple factors? The fundamental difficulty of social program evaluation stems from the fact that we cannot observe what would have happened without a program, in other words, the counterfactual. Therefore, one possible approach for evaluating regional economic development strategy is comparing the outcome changes with a similar region without the policy, but it is hard to find a valid comparable region having *all* possible relevant factors same as the region with policy treatment. Indeed, economic development literatures frequently mentioned the misspecification and measurement issues due to following reasons: 1) relatively long time periods are expected to induce economic outcomes from the program, 2) it is hard to separate the partial effect solely resulted from each individual program, and 3) it is hard to measure the expected outcome if qualitative aspects are involved.[14]

Second, because state R&D tax credits vary and their effects depend heavily on state tax structures, these differences make it quite complicated to compare and evaluate them. State R&D tax credit programs have been developed under each state's own needs and therefore, each of them is unique and distinguished. Hall and Wosinska (1999) mentioned that comparative studies on evaluating the state R&D tax credit policy would be time-consuming for collecting historical data of individual state tax system.

[14] For the detail, see, for example, Reese and Fasenfest (1999) and Bartik and Bingham (1997).

Third, as an important R&D feature, spillover effects should be considered for better evaluation. As a type of public good, R&D is featured as being at least partially non-rivalous and non-excludable, which means that one individual's (or firm's) utilization of R&D does not preclude its utilization of other individuals (see, for example, Arrow, 1962a; Romer, 1990; Spence, 1984). However, it is also generally believed that the benefits from knowledge spillovers are not entirely free, in that the firms should increase their own level of R&D and absorptive capacities for obtaining those benefits.[15] Based on these arguments, the effects of R&D relevant policies do not confine to the firms receiving benefits directly, rather extend to the firms receiving benefits *indirectly* through spillovers.[16]

Fourth, there is a lack of data availability, such as the detailed industry level R&D spending by state. To the contrary, the detailed R&D spending data by industry at the national level (from the NSF) is available.

With the notion of the above four possible difficulties involved in evaluating state R&D tax credit programs, as regional level economic development policies and R&D-relevant policies, the possible evaluation strategies to be utilized in this study are briefly introduced in the next section.

2.2 EVALUATION STRATEGIES IN THIS STUDY

In order to overcome some difficulties mentioned above, this study aims to develop evaluation strategies by applying the quasi-experimental approach, which means using experimentation

[15] For the detail, see, for example, Cohen and Levinthal (1989) and Geroski (1995).

[16] For further issues relating to spillover effects, see Klette, Møen, and Griliches (2000).

without randomness.[17] In particular, the non-equivalent control group design is measuring an outcome change by comparing the treated and the untreated groups, which are selected with non-randomness. It is appropriate for this study because some states have R&D tax credits and some states do not, and they are not randomly distributed. In addition, this study uses the interrupted time-series design, which measures an outcome change before and after treatment by using a relatively large number of observations for same variable over time. This is also applicable in this study based on the availability of multiple historical observations of economic outcomes over time. These two possible quasi-experimental designs can be combined as well.

The major and primary hypothesis to be tested in this study is whether state R&D tax credits have positive effects on state economic development, in terms of increasing private R&D activities such as R&D spending and employment. Under the quasi-experimental design chosen above, the above hypothesis could be modified as following: whether *the outcome changes of states having tax credits before and after operating programs* are larger than *those of states not having tax credits.* In this setting, the difference between two groups indicates the pure outcome change from the program because we assume the outcome change of the latter group captures the possible outcome change from other relevant factors except program.

However, what we have is only a treatment group and a control group that are non-randomly distributed and therefore not equivalent in a quasi-experimental setting. This non-randomness does not ensure a causal relationship of a treatment and its effect because other possible explanations can be involved. In other words, two comparison groups are quite

[17] This terminology was introduced by Campbell and Stanley (1963), who are pioneers to develop the quasi-experimental approach for evaluating social programs. For the detail, see Campbell and Stanley (1963), Cook and Campbell (1979), and Shadish, Cook and Campbell (2003). Randomized social experimentation can be an alternative, but it is mostly not applicable for social settings.

heterogeneous not only by the program but also by other factors. For dealing with this heterogeneity, recognizing other possible explanations, constructing a valid counterfactual, and reducing systematic differences of two comparison groups as much as possible are key for this study.

Previous studies have used similar methods or similar programs for evaluating regional economic development policy. These include enterprise zone analyses (i.e. Dowall, 1996; Greenbaum and Engberg, 1998; Bondonio and Engberg, 2000) and a study on the Small Business Innovation Research (SBIR) program and science parks (i.e. Wallsten, 2001). The present study differs from these studies because the regional effects of R&D tax credit policies are not the same as those of 'area development' policies.[18] In general, most state R&D tax credit programs cover every qualified R&D expenditure, regardless of location in the state, and thus are not limited to a specific geographical location in a state, such as a particular zone or park.

To deal with this difference of policy itself, two major evaluation strategies are suggested. The first set of strategies is deciding the appropriate methodology. One of them is applying the difference-in-differences (DD) method and the difference-in-difference-in-differences method (DDD method). The DD method is conducted for measuring policy effect by comparing outcome changes (defined as before and after program) of two comparison groups which are assumed to be similar at least with their pre-treatment outcome trends. As the advanced form of the DD method, the DDD method allows us to measure this policy effect more precisely by comparing *outcome change of influential group within a treatment group* with *those within a control group* by separately constructing influential and non-influential groups within each two comparison group, not as a whole. In addition, in order to control the systematic

[18] Even though SBIR is not a region-based policy, the analysis is performed by the regional level, which called as SBIR-county.

difference of two groups, the covariates, which are observable and can explain the change of outcome variables, are added within each specification and by doing so the explanatory power can increase.[19] Indeed, by using multiple observations of outcome variables over time and controlling time-specific effects and cross-sectional fixed-effects, which is inspired by the interrupted time-series design, we can also control the possible effects from unobservable variables.

Next, this study adopts the matching method incorporating the DD method. Valid control groups will be selected by finding a pair from observable variables and estimates will be more precise by using the difference-in-differences estimator which can eliminate possible effects from unobservable variables. However, a relatively large number of observations are required for conducting the matching method, and therefore this method will be applied to the firm level analysis only. In sum, both the DD/DDD methods with covariates and the matching methods incorporating the DD method could estimate the policy effect precisely by controlling the possible relevant effects from both observable variables and unobservable variables.

Together with these analytical strategies, this study also aims to investigate the effectiveness of the tax credit by using (1) the multiple estimates which are obtained from multiple methods (the DD method and the matching method); multiple outcome variables (R&D spending and employment); two industrial level analyses (all industry and high-technology industry); and level and log analyses; and (2) multiple datasets which are constructed at the state level and at the firm level. This is the second set of strategies in order to provide more reliable empirical evidences, while most of previous studies provided the single estimate from the single dataset.

[19] This method is also called as the regression-adjusted DD method (Meyer, 1995).

As the primary effect of the state R&D tax credits, I compare average changes of R&D spending before and after the operation of the program. Then, this analysis becomes more informative by estimating possible relevant and general economic outcomes such as employment. Then, I use both the state-level and the firm-level data to develop rich empirical evidence. Indeed, within both levels, the all industry group and the high-technology industry group are examined separately for providing another important evidence of policy impact by the industrial category. This break-down also allows doing the DDD analysis. Finally, the level and the log analyses provide the different interpretation of the policy effect in terms of the actual change of R&D spending and employment and the percentage change respectively. By doing so, this study can deal with the long-lasting question that is whether there is any outcome change purely from the policy program, not combined with other possible changes from other factors.

Next, as one of available datasets, the firm level data allows many variations of analyses that are not applicable to the state-level data. First, based on the large number of observations, the matching method can be applied and indeed within this analysis the county-level data can be included for finding better matched pairs. By constructing counterfactuals at the firm level, different firm characteristics can be controlled. In addition, the firm-level data allows testing for a sub-hypothesis related to firm size: are there any different effects of the state R&D tax credit depending on firm size? This is one of most interesting and significant questions in the R&D-related literatures. It is mostly believed that R&D activities differ by firm size and therefore these differences result in different contributions on the growth of R&D and the subsequent economic impacts.[20] Related to this issue, the R&D tax credits are also expected to generate different effects depending on firm size based on their utilization rules. In detail, the calculation

[20] This argument stems from Schumpeter (1934, 1942) and later, Galbraith (1952) and Arrow (1962b) developed this idea further.

method as an incremental way and the necessity of taxable income as a prerequisite could result in different effects to different-sized firms. Therefore, this study also aims to find an empirical evidence of these possible differences. This hypothesis is tested separately by estimating outcome changes of three different sizes of firms (i.e. small, medium, and large firms) and then comparing each effect. Firm size is defined by the number of employees as indicated at Table 1.

Table 1. The definition of firm size by number of employees

Definition of firm size	Number of employees
small firms	less than 200
medium firms	between 200 and 500
large firms	more than 500

In this chapter, I examined the difficulties involved in evaluating state R&D tax credit programs, as one of regional economic development policies and one of R&D-related policies. These difficulties include the difficulty of measuring effects of individual regional economic development policy within the regional contexts, the differences in state R&D tax credit programs by state, the necessity of including spillover effects for effects of R&D-relevant policies, and data availability. Based on these contexts, I briefly introduced major strategies for evaluating state R&D tax credit programs in this study. These strategies include developing analytical methods and utilizing multiple estimates and datasets for providing the more reliable and strong evidence of the program effect.

In the next chapter, I explore the relevant literatures. They are broadly categorized by R&D-related theories and regional economic development theories. The first set of theories tell us why R&D is important, to what extent R&D is performed and linked with economic growth, and why government intervenes private R&D and so on. The second set of theories show us to what extent private R&D can play a different role of promoting regional economic development within a different perspective. Then important factors for regional economic development are explored and these factors are useful for constructing rival hypotheses and selecting either covariates or matching variables within empirical specification of this study.

3.0 THEORETICAL UNDERPINNINGS

In the first part of this chapter, I examine the theories that focus on the role of R&D and technological progress for explaining economic growth. These theories are categorized by differences between endogenous growth theorists and Schumpeterian approaches.

Endogenous growth theorists explain that the economy grows with the accumulation of R&D or human capital, which are understood as being similar to conventional inputs with some unique aspects within neoclassical equilibrium model. Schumpeterian theorists emphasize the role of entrepreneurs and innovations in economic growth by applying the concept of "creative deconstruction," in which new economic systems are continuously created to replace the old ones.

Then, based on these theories, the property of R&D with the linkage of government intervention is investigated in detail. R&D is characterized as being non-rivalous and non-excludable and generating positive externalities through spillovers. These characteristics provide the legitimization of government intervention based on the possibility of underinvestment in the private market mechanism but the possibility of generating the higher social rate of returns than the private rate of returns.

In the second part of this chapter, I discuss the theories that explain economic growth at the regional level. These theories include economic base theory, location theory, product/profit cycle theory, and flexible production theory. In particular, each theory is explored focusing on its theoretical basis supporting the R&D-related industry or the high-technology industry. Then, I enumerate significant factors for regional economic growth for understanding the overall environment affecting regional economic growth, which provides the important basis of building relationships between R&D policies and their outcomes.

3.1 THE NATURE OF R&D

In this section, I examine the nature of R&D. First, two important theories focusing on the role of technological progress in economic growth are introduced. These are endogenous growth theorists and Schumpeterian approach. Then, based on these theories, I discuss the characteristics of R&D and their linkages with government intervention, in detail.

3.1.1 Endogenous growth theorist approach

Endogenous growth theory, as a modification of neoclassical economic growth model, explains that technological progress promotes economic growth. Within the neoclassical economic growth model, the emphasis on the role of R&D can be traced from Solow (1957), but he indicated the importance of technological change only as an unexplained part.[21] Thereafter, endogenous growth theorists including Romer (1986, 1990) and Lucas (1988) have developed theoretical linkages of technological innovation with economic growth as an additional input of production. In detail, additional inputs for innovations, such as knowledge and human capital,

[21] As following studies, Stoneman (1987) found the high correlation of the Solow's residual with U.S. growth and Denison (1967) argued the residual explains 57% of productivity growth.

can be accumulated by education, training, research, on-the-job-training, process innovation, and product innovation (Aghion and Howitt, 1992). Then, accumulation of these additional inputs determines the growth by increasing productivity and more importantly, generating increasing returns to scale (Romer, 1986). The possibility of generating increasing returns to scale is an important and distinguishable aspect of knowledge and human capital compared with conventional inputs such as capital and labor.[22] In this way, endogenous growth theory can explain economic growth beyond the neoclassical economic growth model in which the factor accumulation generates diminishing returns to scale and productivity growth is treated as an unexplained part.

The major sources of economic growth, which are knowledge and human capital, generate externalities through their sharable characteristics. Namely, knowledge could be utilized to invent multiple products and skilled workers could positively affect other workers without losing their own skills. Importantly, economic growth is accomplished with the accumulation of these inputs by increasing productivity as an unintended and unrewarded ways through externalities or spillovers.

Some endogenous growth theorists have emphasized the role of R&D and knowledge as a primary source of economic growth (see, for example, Romer, 1990; Grossman and Helpman, 1991a; Aghion and Howitt, 1992), while others have emphasized human capital (skill) accumulation (see, for exmaple, Lucas, 1988; Jones and Manuelli, 1990; Rebelo, 1991).[23]

---

[22] It is generally agreed that doubling labor and capital does not result in doubling outputs, rather increasing outputs in less-than-double amount. However, in case of knowledge and human capital, doubling these inputs could result in more-than-doubling outputs through increasing productivity.

[23] For comparison of these two theoretical divisions as well as empirical evidences, see Klenow (1998). About their different policy suggestions, see Romer (1993). Bucci (2002) distinguished these theories as “Research and Development (R&D)-based model” and “AK-based model” respectively.

Following the first line, the increases in R&D result in new idea developments and ultimately, new product developments, which are understood as technological progress. Importantly new ideas can be shared as one of nonrival goods, in other words, same ideas can be involved in developing multiple capital and intermediate goods for increasing productivity through spillovers (Romer, 1990). Technological progress is the fundamental source of economic growth and therefore the policies emphasizing the role of university, new idea commercialization, and intellectual property rights are recommended (Klenow, 1998).

Within the second line of the endogenous growth theory stressing the role of human capital, the learning-by-doing process is essential for economic growth. The skilled worker could increase his own productivity and at the same time encourage other workers increase their productivities through spillovers. Therefore, policies encouraging human capital accumulation, such as job training and education, are recommended from this line (Klenow, 1998).

As an effort to combine these two theoretical lines, human capital and technological progress are considered as having complementary relationships.[24] In detail, the productivity depends on the aggregate level of skills and technological progress is possible from new ideas generated by high skilled workers.

Although the endogenous growth theory provides the better explanation of long-run economic growth based on the more realistic assumptions than the neoclassical approach, the difficulty of measurement for empirical testing is involved. Namely, it is hard to measure R&D, knowledge, and its spillovers. Therefore, some efforts to create proxies of them have been made. R&D has been measured by using number of patents, R&D expenditures, innovation counts, and so on (see, for example, Crosby, 2000; Griliches, 1992; Jaffe, 1986). Also, spillovers have been

[24] For work that integrates these two lines of research, see Arnold (1998) and Blackburn, Hung, and Pozzolo (2000).

measured by using geographical distance, industrial distance, or patent citations (see, for example, Anselin, Varga and Acs, 1997; Caballero and Jaffe, 1993; Feldman, 1994; Griliches, 1992; Jaffe, Trajtenberg and Henderson, 1993).

Endogenous growth theory provides an important theoretical basis for explaining long-run economic growth with possibility of increasing returns to scale. For this, technological progress is considered as the most important factor and being accomplished by the accumulation of R&D and human capital. R&D and human capital is understood as the primary source of increasing productivity and generating positive externalities, which yields continuous economic growth without decreasing returns to scale.

In the next section, I explore the importance of R&D and human capital for economic growth under Schumpeterian approach, in which technological progress is the most important factor for economic growth as like endogenous growth theorist approaches, however the process of reconstructing new economic system is emphasized on. This theory includes the discussion of different roles of small firms and large firms in innovation and economic growth.

3.1.2 Schumpeterian approach

Schumpeter (1934, 1942) long ago provided another theoretical explanation of innovation and economic growth. He argued that economic growth comes from incessant innovations thorough entrepreneurship, rather than the neoclassical static equilibrium model of demands and supplies. Innovative activities, which Schumpeter names as "technological progress," include inventing new products, developing new production processes and transportation systems, creating new markets and industrial organizations (Scherer, 1992). These are able to be accomplished by entrepreneurs who possess new ideas within an expectation of available opportunities and appropriate rewards, and establish new and independent firms. Then, increasing market demand results in economic growth, and newly developed market structures replace old ones with redistributing wealth. This process is called "creative destruction" (Schumpeter, 1942, p.83).[25] Schumpeter (1942) described the process of "creative destruction" as "a process of industrial mutation that incessantly revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one" (p. 83), as an essential dynamic of capitalist system.

Importantly, entrepreneurs play an essential role for economic growth, and they are characterized as outstanding individuals/groups enthusiastically introducing new products/ systems/process with creative leaderships under risk and uncertainty, and consequently leading to heroic innovations. Therefore, within this theory, the market demand is not given from consumers, but created by entrepreneurs. Schumpeter (1934) also mentioned economic growth takes the form of a cyclical process due to uneven distribution of innovation occurring spontaneously and intermittently. In this sense, the Schumpeterian growth theory overemphasized the role of outstanding type of innovation from new knowledge but underemphasized the role of steady and incremental technological progress from the stock of previous available knowledge. Mokyr (1990) insisted that the spectacular break-through innovations and relatively small technical improvements are complementary and the latter is as essential for economic growth as the former.

Schumpeter also raised another important issue regarding to the relationship between firm sizes and R&D activities. Within his theory there are two different notions related to this issue. Namely, while he argued that small firms are more innovating in his earlier work (1934),

---

[25] Refer to Scherer (1992) for following developments of Schumpeter's theory and related empirical works.

he (1942) later explained that large firms are more appropriate for innovation due to their monopolistic power and large amount of available finance sources (Scherer, 1992). Earlier small and new firms formatted by entrepreneurs were considered as playing a central role in economic growth and later large firms were recognized as driving innovation in reality.

This relationship between firm size and R&D activities has continued to be an important and long-debated issue within R&D-related literatures. Galbraith (1952) argued large firms could utilize R&D expenditure less riskily and more efficiently than small firms with the emphasis of the "costness" of R&D, such as the necessity of considerable time and risk. Arrow (1962b) maintained that large firms could take advantage to capture the benefits through the property rights, which provides the higher incentive. In sum, it is hypothesized that R&D activities are likely to be skewed to large firms based on market imperfections, scale economies, and complementarities between R&D and other activities (see, for example, Lee and Sung, 2005 and Cohen and Levin, 1989).

Related to this issue, some empirical studies focused on finding the direct relationship of R&D expenditure and firm size, in other words, whether there is any proportional increase of R&D expenditure by firm size based on the advantage of large firms for innovative activities. The results of these studies are somewhat mixed. Link (1980) and Pavitt (1983) found that there is a positive, linear relationship. Meanwhile, Acs and Audretsch (1991) and Bound, Cummings, Griliches, Hall, and Jaffe (1984) argued that there is a non-linear, inverted U-shaped relationship, in other words, less-than-proportional relationship. The less-than-proportional relationship implies that as firm size increases, R&D expenditures increase at a certain point, and then decrease. Cohen, Levin, and Mowery (1989) found neither clear nor significant relationship between firm size and R&D expenditure with controlling industry effects. Recently Lee and

Sung (2005) found that this relationship depends on firm-specific technological competences that are defined as absorptive and learning capabilities and measured by the elasticity of the productivity of R&D with respect to R&D expenditure. They found that a less-than-proportional relationship is generally observed, however, firms having high level of technological competences show a more-than-proportional relationship.

In addition, some empirical studies focused on finding the relationship between innovative output and firm size, in other words, the productivity of R&D by firm size. These studies showed the similar result which tells us that the productivity of R&D tends to be decreased by firm size. Audretsch (1991, 1995) indicated small firms could take advantage through their higher ability of innovation and Chakrabarti (1991) also found an evidence of small firms' higher rate of innovation. Acs and Audretsch (1991), Bound, Cummings, Griliches, Hall, and Jaffe (1984) and Hausman, Hall, and Griliches (1984) suggested that small firms are more productive and innovative based on the higher number of patents per spending.[26]

Schumpeterian growth model explains that economic growth is accomplished by the entrepreneurs' continuous innovative activities including new products, new production processes, and new organization, subsequently resulting in continuous changes of rebuilding new market system with demolishing old ones. In addition, it is theoretically believed that large firms take an advantage in innovation due to scale of economies and market imperfections to overcome high risk and uncertainty. However, empirical findings show some mixed results about the relationship between firm size and R&D expenditure as shown above.

---

[26] However, some questions arised about these findings such as the measurement issue of R&D expenditure (for example, Cohen and Levin, 1989). There is common belief of biased R&D expenditure reporting, which is small firms usually under-report their R&D expenditures.

This section and the previous section tell us that R&D and innovation are fundamental for economic growth. Based on these theories, I discuss the nature of R&D and the rationales for government intervention to private R&D in the next section. These discussions include findings of relevant empirical analyses and possible policy impact at the regional or state levels.

3.1.3 Legitimization of government intervention

In this section, I examine the properties of R&D and their relationships with government intervention in detail. Theoretically, R&D can be understood as one of public goods, as at least partially non-excludable and non-rivalrous, which means that one individual's (or firm's) utilization of R&D does not preclude its utilization of other individuals (or firms) at least partially (see, for example, Arrow, 1962a; Romer, 1990; Spence, 1984). Accordingly, R&D tends to be underinvested for avoiding its common utilization without proper compensation and therefore, it is difficult to reach the optimal point of R&D investment only within a private market mechanism, which is recognized as one of market failures. This provides the rationale for government intervention to private R&D in order to increase R&D investment up to the optimal point, or at least close to the optimal point. Government could increase R&D investment either by spending its own funding and performing its own R&D as a form of direct intervention or by encouraging the private R&D activity through a variety policy tools as indirect intervention.[27] The R&D tax credit is one type of indirect government intervention.

[27] For the comparison on these alternatives, refer to Hall and Van Reenen (2000). Office of Technology Assessment (1985) pointed out the different roles of indirect and direct government funding and therefore argued the necessity of their complementary utilizations.

However, the legitimization of policy intervention, including the R&D tax credit, is somewhat controversial. There is a possibility of windfall gains, indicating that the increase in private R&D activities could be achieved without credits. Over-subsidy, which means providing funds or resources beyond the necessary level, is one type of windfall profit from government intervention. To deal with this possible problem, the appropriateness of program should be evaluated by estimating the pure effect of program, which is only correctly recognized by comparing the situation with outcomes from no intervention.

As a different point-of-view, there is another theoretical possibility in a neoclassical model, which is overinvestment, namely the investment over the optimal point for a society. Because it is hard to know where the optimal point is, it is also hard to know whether government intervention increases the R&D investment up to the optimal point or over the optimal point. However, this issue could be ignored because previous work revealed that the gap of private and social returns is substantially high, and therefore overinvestment is rare in real world (Hall and Van Reenen, 2000). The other controversial issue is government capacity for efficient intervention, which is more related to direct funding. In addition, there are mixed results from empirical evaluation studies on government intervention for stimulating private R&D (see, for example, Bartelsman, 1990; Griliches, 1986; Nadiri and Mamuneas, 1994). With these on-going controversial issues involved in government intervention, the federal-level R&D tax credit has been renewed eleven times as a temporary law for twenty-four years and expired again in the end of 2005. At the state level, by 2003, some state R&D tax credits have already expired, operated as temporary, or utilized in only restricted ways.[28] Of these, eighteen state laws were permanent while twenty state laws were temporary.

As the other property of R&D as one of public goods, the social rate of return of R&D investment is much higher than its private rate of return due to its externality, which means that R&D is sharable by multiple actors without additional costs (see, for example, Griliches, 1979, 1992; Romer, 1986, 1990; Lucas, 1988). Through spillovers, one firm's innovations and improvements can positively affect other firms' productivity without appropriate compensation. Through spillovers, one person's skills and knowledge can positively affect other person's skills and knowledge in unintended or unrewarded ways. Namely, the externality, also called as spillovers or knowledge spillovers, is understood as the primary source of increasing productivity and continuously positive economic growth.

Empirically there are plenty of studies pointing out that R&D investments have a positive effect on productivity, supporting R&D investments as an important ingredient for sustainable economic growth (see, for example, Griliches, 1995; Mowery and Rosenberg, 1989; Nadiri, 1993; Nelson, 2000). Griliches (1992) concluded that R&D spillovers exist and generate large effects, resulting in the higher social rate of return than the private rate of return by reviewing previous studies. Jones and Williams (1998) argued that the R&D is substantially underinvested based on the R&D-based growth theory and previous findings which told us that the much higher social rate of return to R&D investment (ranged from 30% to 100%) than the private rate of return (average 7%).

[28] For example, Michigan R&D tax credit can be applied to only pharmaceutical industry and Texas has a temporary R&D credit law. These states are R&D intensive states within the U.S. based on amount of private R&D investment.

The important feature of knowledge spillovers is that they are geographically localized. Namely, it is generally believed that some geographical barriers exist for spillovers and this is why individual firms and people are eagerly located near other firms and people who possess the advanced knowledge and skills. Emphasizing on proximity and the face-to-face communications for learning from others (see, for example, Lucas, 1993 and Glaeser, 1999), supports the possible geographical barrier for knowledge spillovers. Von Hipple (1994) also supported this argument by distinguishing knowledge and information and especially introducing sticky knowledge that is highly contextual and uncertain and therefore can be transmitted via fact-to-face interaction.

Jaffe, Trajtenberg, and Henderson (1993) found an empirical evidence of geographically bounded knowledge spillovers by using the patent citations. They argued that the closeness is fundamental for create new patents from the previous patents, even though patents are represented as a type of knowledge easily accessible anywhere by collecting and filing with highly well-organized system. Almeida and Kogut (1994, 1999) also found empirical evidences of localized knowledge spillovers by using the similar data with Jaffe, Trajtenberg, and Henderson (1993), namely through geographical location of patent holders and patent citations. Audretsch and Feldman (1996) found that there are some efforts to create clusters for gaining benefits of these spillovers especially among R&D intensive industries. They emphasized that the geographical distance matters for knowledge spillovers, in particular, spillovers for tacit knowledge, and supported this argument based on the highly significant relationship between the concentration of economic activities and the concentration of knowledge generators that are measured by industry R&D, skilled worker, and university research. Even though there is a study which argued that the geographical distance is less important for knowledge spillover and rather it depends on specific needs and prompt searches of private sectors (Mowery and

Ziedonis, 2001), there are much larger volume of studies supporting the geographically localized knowledge spillovers, especially spillovers for sticky knowledge or tacit knowledge.

In sum, the geographically localized feature of knowledge spillovers is supported theoretically and empirically as shown above. It tells us that the benefits from knowledge spillovers, as a major source of increasing social rate of returns, tend to be geographically bounded. Therefore, promoting knowledge spillovers for economic growth should be one of fundamental goals of the regional level or the state level regional economic development policies. More importantly, this geographically localized characteristic of knowledge spillovers provides a fundamental rationale for the state level or the regional level policies supporting private R&D, including state R&D tax credit.

The benefits from knowledge spillovers can vary depending on region's capacity of supporting spillovers. If some regions establish more developed regional economic networks in order to mobilize and utilize regional knowledge more efficiently, they could enjoy much larger benefits from externalities and vice versa. It implies that regions having a similar level of knowledge could accomplish different levels of economic growth by their different levels of knowledge spillover. Moreover, if some regions have substantially large amounts of R&D investments and also well-established network systems but other regions do not, the disparity between prosperous regions and backward regions gets larger and deeper. This argument provides another theoretical basis of the importance of government intervention to private R&D. In this sense, the development of overall R&D-friendly environment within a state is one of important tasks of state governments.

The spillover effects also provide the possibility of indirect effects on firms that don't receive tax credits. The possibility of indirect effects also stems from an argument that tells us individual firms are expected to make their own investments for obtaining benefits from spillovers. Cohen and Levinthal (1989) elaborated this argument by indicating the necessity of increasing the firms' absorptive capacities to utilize the other firms' R&D. Geroski (1995) supported this argument as firms should invest their own R&D as the efforts to learn from other firms. Klette, Møen and Griliches (2000) also insisted that the benefits from the spillovers are not completely free, and rather complementary efforts are required.[29] Accordingly, these catch-up efforts could be indirect effects of the tax credit.

Related to this issue, there is an analysis which tells us that most knowledge spillovers occur from large firms and universities to small firms within geographical proximity (Acs, Audretsch, and Feldman, 1994). They used knowledge production function for finding the effect of firm's own R&D and university R&D (as a proxy of spillover) on number of innovation (as a proxy of innovative output) by firm size and they maintained that spillover is a major source of innovative activities of small firms, whereas large firms do their innovations by using their investments.

There is the other analysis which tells us the most beneficiary firms from the state R&D tax credits are large firms (Sigalla and Viard, 1999). OTA (1985) also mentioned the majority of beneficiary of the federal R&D tax credit are large firms.

These two kinds of analyses also support the legitimization of state R&D tax credit in the sense that even though large firms are major and direct beneficiaries of the state R&D tax credit program, small firms are possible indirect beneficiaries through knowledge spillovers as well. In other words, the benefit of this program is expected to be broadly applicable from large firms to

[29] For empirical evidences, refer to Cohen and Levinthal (1989) and Branstetter and Sakakibara (1998) and Geroski, Machin and Van Reenen (1993).

small firms and therefore every R&D-intensive firm should be a possible beneficiary in either a direct or an indirect ways.

Until now, the significance of private R&D was explored with a perspective of the nature of R&D. It can be summarized as follows. First, R&D, as a type of public goods, generates positive externalities but is likely to be underinvested in the private sector. Second, based on its positive externalities, in other words, knowledge spillovers, it is expected that R&D generates the higher social rate of return. Third, knowledge spillovers tend to be geographically localized. These arguments are strongly supported by some empirical analyses. These natures of R&D provide important rationales of government intervention to private R&D. In particular, the localized feature of knowledge spillovers supports the regional level or the state level policies promoting private R&D.

In the next section, I discuss the role of private R&D under several perspectives of regional economic development theories. Each theory is examined focusing on the role of the R&D-intensive industry or the high-technology industry, in the sense that this type of industry is an industry performing the higher R&D activities and is recognized as playing a primary role in economic growth. Then, I examine important factors determining regional economic growth for understanding plausible explanations for possible economic outcomes from state R&D tax credit programs.

3.2 REGIONAL ECONOMIC DEVELOPMENT THEORIES

Regional economic development theories are broadly characterized as follows: (1) economic base theory, (2) location theory, (3) product/profit cycle theory, and (4) flexible production theory.[30] These theories provide rationales for supporting R&D investment for economic growth especially at the regional level from different perspectives. These theories are different in their emphases on shaping the regional economy and the policy implications. Indeed, the role of the R&D-intensive industry, or the high-technology industry for regional economic growth is also distinguished by each regional economic development theory. Then, the significant factors affecting regional economic growth are reviewed based on these regional-level theories. By doing so, the overall environment surrounding R&D-relevant policies and their outcomes can be constructed and recognizing this overall environment is essential for identifying rival hypotheses in this study and important variables that should be included in empirical specifications.

3.2.1 Economic base theory

First, economic base theory assumes that the regional economy consists of two sectors, which are the basic sector and the non-basic sector. The basic sector, in other words the export sector, produces exportable goods based on demand to the external region and carries out economic activities related to non-local residents, while the non-basic sector, or local sector, makes products based on the within-region demands and performs economic activities related to local residents. This theory, therefore, explains that the regional economy grows primarily from the increase of the basic or export sector rather than from the increase of the non-basic or local

[30] For overall assessments of these theories, refer to Blakely (1994, Ch 3), Blakely and Bradshaw (2002, Ch 3), Bingham and Mier (eds.) (1993, Ch 1, Ch2, and Ch7) and Malizia and Feser (1999).

sector. Here the regional output depends on the size of external demand for exports, and the increase of the non-basic sector is subsequent to the increase of the basic sector.[31]

Economic base theory suggests the possibility of regional economic growth from appropriate policy interventions, for example, attracting industries through incentives and subsidies.[32] It is assumed that regional competitiveness could be improved through intentionally produced environments together with natural endowments. Blakely (1994) maintained that current policies supporting entrepreneurial and high-technology industries are based on this theory in the sense that these industries generate the higher multiplier effects than local service sectors.

Economic base theory provides the sources of regional growth and the corresponding policy implications from its simplicity. However, it ignores the importance of local service industries and the balance of whole regional economy. Indeed, this theory is more appropriate to explain the short-term economic growth, but inadequate to understand the long-term growth because a local economic structure, including interactions of the export sector and the local sector, has been changed over time (Malazia and Feser, 1999).

Within economic base theory, which explains the regional economy can grow by supporting the basic or export sector, the high-technology industry, or the R&D-intensive industry plays an important role in regional economic growth as one of the export-based industries. In the next section, the arguments are moved to significant determinants of the high-

[31] It is assumed that the supply could be generated by the demand force. However, in the neoclassical model, rather the supply is considered as a constraint and therefore the increase of the supply result in the increase of the demand based on Say's law.

[32] For example, the City of Chicago and the State of Illinois had offered the Boeing Corporation a generous package of tax incentives and other subsidies and the Boeing Corporation's headquarters were relocated at Chicago.

technology industry location within location theory. The arguments relevant to the role of taxes for decision making for firm and industry location are included.

3.2.2 Location theory

Location theory has been developed to shed light on what are influential factors for deciding location of economic activities including firm and industry. Through recognizing and analyzing these factors, this theory can provide important policy implications for regional economic growth. In particular, by diversifying locational factors, the explanations for the high-technology industry location have been developed.

The importance of location was paid attention long ago and classical location theory or Weberian location theory (originally developed by Weber, 1909) focused on finding optimal location between raw materials and final markets by minimizing total costs, including transportation costs, accessibility of materials and markets, and availability of labor. Transportation costs were the primary concern in this theory (see, for example, Thisse, 1987) and mostly existing spatial patterns were analyzed (see, for example, Christaller, 1933; Hoover, 1948; Lösch, 1954). Markusen, Hall, and Glasmeier (1986) insisted that the classical location theory is less relevant to high-technology industry because this type of industry is featured by a low demand of raw materials and a low value-to-weight ratio.

Then, location theory was enriched by incorporating multiple factors, such as skilled labor force, market size, governments, technological capabilities, education, agglomeration economies, and the quality of life. In detail, Isard (1956) maintained that agglomeration economies are one of important factors for firm location. Schmenner (1982) provided broad empirical studies of locational factors and he found out that firms' need to maintain labor force motivated to keep their original place. He also found the right to work laws made most efficient effects on firm locations among the various policies. Blair and Premus (1987) enumerated important locational factors as productivity, education, taxes, community attitudes and the quality of life along with traditional factors.

Under location theory, taxes are considered as one of important influential factors for firm location. Even though there is a consensus that tells us that the taxes do not have a strong effect on new firm births (see, for example, Carlton, 1983; Blair and Premus, 1987), Bartik (1991) insisted that there is a potential influence of local incentives in the situation of that expected influences of other factors are similar among final alternative locations.

Recent location theory can provide the better policy implications in the sense that multiple influential factors for deciding location can be understood as being created and developed within a region through policies rather than just pre-determined regional characteristics. These factors also include some specific factors which are closely relevant to the high-technology industry. There are several empirical studies looking for important locational factors especially for the high-technology industry. For example, Schmenner (1982) maintained that the availability and the cost of the skilled labor with the proximity of university is the most important influential. Myers (1987a, 1987b) found that the quality of life is particularly important for the high-technology industry.

Location theory provides the important basis for regional economic development strategies especially related to the high-technology industry, because the "location" gains the more important meaning in the era of information and technology society than any other previous period. Namely, even though location seems to be meaningless with well-developed transportation and communication technology, some (see, for example, Almeida and Kogut, 1999; Jaffe, Trajtenberg and Henderson, 1993; Anselin, Varga and Acs, 1997) argued and proved that the proximity and frequent contacts are getting more important for accomplishing technological innovation through knowledge spillover, and eventually economic growth. In the next section, I examine the importance of R&D and the high-technology industry under product cycle theory emphasizing their roles in developing products within a cyclical way.

3.2.3 Product cycle theory/Profit cycle theory

Product cycle theory, first developed by Vernon (1966),[33] models the process of regional economic growth focusing on innovation. This theory is grounded neoclassical economics, using a partial equilibrium framework and specialization in trade (Malizia and Feser, 1999). Within product cycle theory, regional economic growth is accomplished by recurring cycles of creating new products, maturing products and diffusing them to other regions, and standardizing products. In detail, a newly invented product tends to originate in a more-developed region that is characterized as possessing an advantage based on its advanced knowledge, the more efficient market, and diverse demands of high income residents. Then, as the new product matures, which means finding a right market, setting up a right price, and adopting mass production system (Goldstein and Lugar, 1997), the more-developed regions could keep expanding its outputs through exporting products to other less-developed regions. After a certain time period, those products are standardized and less-developed regions become more competitive to produce

[33] He originally modeled product cycle theory at the international context, not the regional context. His theory provided an alternative explanation of international trade for dealing with the "Leontief paradox", which tells us that the United States, as a capital-intensive country, exported labor-intensive products and imported capital-intensive products, which is the opposite of the Heckscher-Ohlin theory.

standardized items, which means less-developed regions have a comparative advantage based on their cheaper production costs, including labor costs.

In order to obtain new advantage or another advantage before the current advantage disappears, new knowledge linked with product developments is essential. Mostly, the high-technology industry, or the knowledge-intensive industry, plays a primary role in inventing new products, which is an engine for regional economic growth. Importantly, more-developed regions could provide the better environment for innovation, such as the higher opportunity of knowledge spillover, the larger and more diverse market, and the larger agglomeration economies, which also attract the more entrepreneurs and the more high-technology industries within the region. Accordingly, at the initial cycle of products, further innovative activities tend to be achieved within more-developed regions, and then new products are invented, produced and exported. However, as products and markets mature, the existing firms tend to move to less-developed regions for finding the cheaper labor and costs for output production. Then, new product developments follow by new firms within more-developed regions for creating another advantage. In this sense, product cycle theory explains new firm formations and the following relocations depending on the product life cycle.

Product cycle theory explains that new and small firms are mostly engaged in product developments at their initial stage. In order to obtain new ideas and find proper markets, they are located at more-developed regions. Utterback (1994) supported this argument in the ground that knowledge spillovers play the more important role in an early product cycle stage than any other product cycle stages, especially for young and new firms, because new products tend to be developed with high risk and high uncertainty and benefits from other firms through knowledge spillovers can reduce initial product development costs. As firms mature and produce standardized products, lower labor costs are more important and therefore they move into less-developed regions. Blair and Premus (1993) pointed out that some high-technology firms keep looking for new products and focus on innovation, which makes them continuously remaining at the initial product cycle stage, suggesting some high-technology firms might not follow these product cycles.

In sum, as products are newly developed, matured, and standardized, the comparative advantage of production moves from more-developed regions to less-developed regions. Therefore, in order to gain new comparative advantage, the continuous product development is required especially for more-developed regions. In addition, following the product cycle, the firms tend to move from more-developed regions to less-developed regions. Therefore, the new and small firms tend to be located in more-developed regions.

Markusen (1985) developed profit cycle theory as a broader version of product cycle theory. She paid attention to corporate decision-making behaviors depending on various economic factors and the given market demand in the context of globalization. Profit cycle theory focuses on the profitability of production instead of the cyclical development stage of product itself. Sorenson (1997) mentioned that profit cycle theory is a well-organized growth theory combining supply-side market conditions with demand changes of products in a cyclical way. Profit cycle theory also provides the possible growth of employment, output, investment, occupational composition, and geographical concentration with a linkage of each profit cycle.

Instead of product life cycles, which are defined as new, maturing and standardized products, profit cycle theory defines profit growth cycles as “zero profit” at the experimental stage, “super-profit” at the innovative stage, “normal profit” at the competitive stage, “normal-plus and normal-minus profit (declined-profit)” at the stage of standardized products, and “negative profit (rapidly declined-profit)” at the final stage (Markusen, 1985). In detail, in the initial stage the profit goes up due to rapid market expansion and then, the profit goes down as a market matures, which means severe competition exists and products are standardized. Then, the profit could be retained or declined depending on market conditions. In the late stage, products are saturated in the market and the profit more declines and disappears.

Markusen, Hall and Glasmeier (1986) underscored that close proximity for production inputs, especially labor pooling, were influential factors for high-technology firm location. They combined the concept of agglomeration economies with profit cycle theory. Duranton and Puga (2001) developed a general equilibrium model based on product cycle theory. They argued that different types of regional specialization are required according to the life cycle of products. In detail, in the earlier stage of product innovation, more diversified regions provide better environment and in the later stage of product standardization, more specialized regions provide better environment, which results in the relocation of individual firms.

Product cycle theory has some similar aspects with economic base theory, in that these two theories emphasize the role of exports for regional economic growth. However, product cycle theory underlies an emphasis on encouraging innovation through generating and utilizing technological knowledge and diffusion of products instead of simply attracting industries. Product cycle theory also has similar aspects with Schumpeter’s economic growth theory, in that these theories explain economic growth as a process of recurring cycles and also underscore the role of innovation. As a drawback, it is hard to explain the growth of service industries because service industries do not grow with creating new products and distributing them, rather they growth with developing relatively intangible assets (for example, Vernon, 1979). Taylor (1986) also criticized that the market is not likely to be properly developed according to product developments and the definition of the product itself is ambiguous. Taylor (1987) also pointed out that products tend to follow unique and developmental ways such as differentiation and improvement, rather to follow homogenous paths of maturations and standardization. Thus, profit cycle theory elaborates product cycle theory in the broader context of relationships between supply and demand and at the industry level. Sorenson (1997) pointed out that empirical supports for profit cycle theory are relatively few while this theory is widely cited and applied in regional development literatures. Based on this notion, he empirically tested the hypotheses in profit cycle theory and found that the growth of employment and output generally followed with profit cycles while geographical concentration were mostly steady across profit cycles.

Product/profit cycle theories provide important policy implications for new firm formations and their relocations. For example, government policies stimulating the production of standardized products and the creation of new products could generate regional economic growth with different ways. Importantly, the notion of that there are basic necessities of analyzing regional capacities and differentiating strategies based on them is major contribution of this theory.

Supporting R&D is especially important to develop new products especially at the initial stage of product cycle. Indeed, knowledge spillovers are also important to develop new products especially by new and small firms because their own investment capacity is mostly not large enough to perform necessary R&D activities, and therefore utilizing external resources are essential for reducing the costs.

3.2.4 Flexible production theory

Flexible production theory is based on the recently developed economic environment, such as new product development system to cope with severe competition and to satisfy the diverse demand of variety and quality through globalization. As an expansion of location theory, its focus was shifted from traditional manufacturing industries to high-technology industries and producer service industries. These newly emerging industries tend to create new types of spatial organization, which are recognized as “new industrial districts” (Piore and Sable, 1984). These new industrial spaces are featured by intensive inter-firm networks, vertically disintegrated firms, and flexible specialization. Within this theory, specific factors for determining the location of high-technology industries, which are quite different from traditional factors, are examined. These distinct location requirements formulate a re-agglomeration of economic activity, mostly stemming from the outside of firms. Flexible production theory is also linked with product cycle theory in the ground that new industrial spaces are varied highly depending on the type of product and the stage of product development.

Based on the localized production system and the strong social division of labor, Scott (1988) explained the formation of small and medium-sized high-technology firms. Because these small and medium-sized firms do not have enough investment capacities under high risk and uncertainty, these firms tend to heavily rely on the available external resources to reduce production costs. Accordingly the gains from externalities are essential for these small and medium-sized firms. Therefore, these firms are mostly concentrated and clustered for finding the localized pool of technological knowledge and higher skilled workers and highly rely on external networks with close proximity.

Storper and Walker (1989) pointed out the general explanation of firm location based on searching profits or minimizing costs cannot adequately reflect the recent firm formation within peripheral areas and therefore, specific needs depending on industrial characteristics should be more paid attention to. Namely, industries grow by their own trajectories, based on different histories, prices of inputs, technology utilizations, under strong competitions. They also stressed that a region is not just given, but chosen, developed, then specialized and differentiated by particular industrial trajectories and strategies and therefore, new spatial structures keep replacing existing ones and uneven spatial development is getting deeper, instead of trickle-down or equilibrium.

There are some efforts to explain the variation of spatial organization. For example, Glasmeier (1988) categorized the product types as "one-of-a-kind," "customized," and "standardized" and argued that "customized" product developments have higher potential of backward and forward linkages and spin-offs. Markusen (1996) identified five different kinds of spatial structures as "the Marshallian," "the Italian variant to the Marshallian," "the hub-and –spoke," "the satellite industrial platform," and "the state-anchored industrial district." Within this paper, she argued that regional specialization should be understood in the broader context of external linkages rather than focusing on localized structure. She also pointed out the existence of multiple forces and institutions which are embedded in external relationships, for example, relationships with large firms.

Flexible production theory well reflects the recent phenomenon of regional economic growth and also emphasizes the role of small and medium-sized firms on innovation. This theory attempts to explain industrial growths as a cyclical way, which is similar with product cycle theory, and analyze important factors for firm location, which is an expansion of traditional location theory. This theory underscores the necessity of specified and differentiated development strategies for individual industries, especially for the high-technology industry and small and medium-sized firms from the notion that the characteristics of new industries, such as decision-making for firm location and their development paths, are quite different from traditional manufacturing industries.

Flexible production theory emphasizes the role of R&D for developing differentiated and localized products instead of standardized and mass-production-based products. Especially, these differentiated and specialized production systems are recognized at the regional level with the emphasis on intensive inter-firm networks and vertically disintegrated firms. Namely, firms are mostly concentrated and clustered for finding the localized pool of technological knowledge and higher skilled workers and highly rely on external networks with close proximity. In this sense, for regional economic growth, the role of R&D and its spillovers are important.

Until now, I examined four regional economic development theories, which are economic base theory, location theory, product/profit cycle theory, and flexible production theory, in terms of basic assumptions and the mechanisms for accomplishing regional economic growth, and the role of R&D and the high-technology industry within these theories.

In the next section I discuss how specific economic factors affect to regional economic growth. These factors are recognized as being specially important with a linkage of the role of R&D and knowledge in economic growth. The discussions include previous empirical researches for findings the relationship between these factors and regional economic growth.

3.2.5 The significant determinants of urban and regional economic growth

In this section I examine the arguments about important determinants for economic growth, especially at the urban and regional level and with an emphasis of the role of R&D and knolwedge. Glaeser (2000) pointed out that the emergence of endogenous growth theory, which emphasizes the role of new ideas and knowledge, evoked an important implication for urban and regional economic growth because cities reinforce creating new knowledge than any other places. Based on this perspective focusing on the role of knowledge, he recognized five fundamental components for determining urban growth, yet still in debate. They are human capital, diversity/specialization, competition, government policy, and initial economic conditions. These will be explained and examined in the following section, focusing on the mechanism involved in, relevant arguments, empirical studies, and measurement issues, which are summarized in Table 2. These five topics are also interwoven and interdependent. For example, knowledge spillovers, generated by human capital, are also a fundamental source of agglomeration economies, and the arguments of agglomeration economies are closely relevant to the arguments of competition for understanding the role of knowledge spillovers in innovative activities of individual firms.

Table 2. The significant factors for urban economic growth

Factors	Mechanism	Empirical studies	Measurements
Human capital	- Increase in opportunity of creating new knowledge and increasing productivity - Increase in opportunity of knowledge spillovers	Rauch (1993), Glaeser (1994), Glaeser, Scheinkman and Shleifer (1995), Simon (1998), Simon and Nardinelli (2002)	- Year of schooling, work experience - Percentage of high education (high school and college) - Occupational mix - Number of high skilled employment
Diversity vs. Specialization	- Urbanization economies as interactions across industry (Jacobs, 1969) - Localization economies as interactions within industry (MAR: Marshall-Arrow-Romer and Porter, 1990)	Glaeser, Kallal, Scheinkman and Shleifer (1992), Henderson, Kuncoro and Turner (1995), Ciccone and Hall (1996), Ellison and Glaeser (1997), Rosenthal and Strange (2001), Dumais, Ellison and Glaeser (2002),	- Gini coefficients - Herfindahl index - Ellison-Glaeser Index - Employment/population density - Establishment density
Competition	- Providing incentives for improving and encouraging innovation through spillovers - Making a barrier and discouraging innovation due to the possibility of spillovers	Glaeser, Kallal, Scheinkman and Shleifer (1992), Feldman and Audretsch (1999)	- Lerner index - Herfindahl index - Population/employment density - Establishment density
Government policy	- Firm location - Wage differentials - Regional outputs	Bartik (1991), Rauch (1995), Wasylenko (1997), Holmes (1998), Saiz (2001), Brown, Kathy and Taylor (2002)	
Initial condition	- Convergence vs. Divergence - Gibrat's law - Zipf's law	Glaeser, Scheinkman and Shleifer (1995), Beeson, Dejong and Troesken (2001)	- Population/Employment - Per capita income

Source: Glaeser (2000), elaborated by author

3.2.5.1 Human capital

First, human capital, in particular skilled workers, is a major source of creating knowledge and increasing productivity for accelerating regional economic growth. Lucas (1988) indicated utilizing existing ideas and creating new ideas is essential for economic growth and these two important forces highly depend on the stock of human capital. Importantly, skilled workers are limited at least in certain times within a region, and therefore their location is substantially meaningful for knowledge utilization and subsequent economic growth. Moreover, their interactions result in significant externalities, namely knowledge spillovers, which lead to the endogenous growth (see, for example, Romer, 1986, 1990; Lucas,1988; Krugman, 1991a; Grossman and Helpman, 1991a, 1991b) and it is widely believed that knowledge spillovers are geographically bounded (see, for example, Jaffe, Trajtenberg and Henderson, 1993; Audretsch and Feldman, 1996). Private firms keep looking for highly educated/experienced persons who possess the-state-of-art idea they need and make a decision for their location to be close with them. People are also trying to obtain the opportunity for learning and improving their ability by face-to-face interactions with them. All of these efforts could result in the regional economic growth.

Glaeser (1999) focused on the learning process from face-to-face interactions and explained that the larger cities could provide the better environment for increasing these interactions based on the density and the proximity. He modeled the level of skills and their variances are a function of city size. Especially young people, who are patient to learn and ready for risk-tasking, could receive much larger benefits from cities and therefore they are willing to stay the cities for learning and imitation. Moreover, the importance of face-to-face interactions would have complementary relationship with telecommunication development. However, the former could not be replaced with the latter and consequently it is predicted that the cities will grow even in the future (Gaspar and Glaeser, 1998, Glaeser, 1999).

There are plenty of findings of the positive effects of human capital accumulation on economic growth by using the metropolitan area data in US. Cities, having more skilled workers and presumably more spillovers, show higher wages and housing rents (Rauch, 1993, measured by average years of education and work experience), higher wage growth (Glaeser, 1994, measured by percentage of college graduates), and faster growth of population, employment, and income (Glaeser, Scheinkman and Shleifer, 1995, measured by percentage of high school graduates). More recently, Simon (1998) and Simon and Nardinelli (2002) showed human capital accumulation contributes to employment growth by using both percentage of college graduates and high school graduates. Simon and Nardinelli (1996) and Mody and Wang (1997) provided an empirical evidence of positive effects on human capital on economic growth in case of English cities between 1861 and 1961, and China in 1980s respectively. Charlot and Duranton (2004) also provided an empirical evidence of economic benefits from communications. They conducted survey recording workplace communication for approximately 6000 French workers, and found there is a positive effect on wage increases from communication externalities and the larger communication in larger and more educated cities.

In sum, based on extensive empirical findings of positive effects of human capital on regional economic growth, human capital, or skilled workers, is a major source of regional economic growth by creating and utilizing knowledge and subsequently increasing productivity.

3.2.5.2 Diversity *vs* specialization

Second, whether diversity or specialization is an impetus for urban and regional economic growth is long-lasting debate in this field (Glaeser, 2000). Originally regional concentration of economic activities was paid attention to for generating positive growth effects by Marshall (1890), who is a pioneer developing the idea of agglomeration economies. He pointed out there are some important reasons why individual firms are located with other similar firms closely. These reasons are 1) labor pooling, 2) accessibility of intermediate/final goods, and 3) knowledge spillovers in order to reduce costs and generate increasing returns. Chinitz (1961) emphasized the role of knowledge spillovers for agglomeration, comparing New York as a highly diversified and fast-growing area and Pittsburgh as a highly specialized and slow-growing area. He underscored the role of entrepreneurship as a primary production factor and it could be encouraged from external economies under competitive inter-industry relationships. With the viewpoint of international trade, Krugman (1991b) stressed the role of the following three factors for regional specialization and concentration: 1) the diversified pool of workers for developing the specialized local division of labor, 2) sufficient customers and suppliers for achieving the economies of scale, and 3) knowledge spillovers.[34] He also mentioned that regional specialization initially formatted by historical events could be lasting through circular and cumulative process of regional economic growth.

Porter (1990) used the concept of comparative advantage for a fundamental basis of regional clustering together with suggesting similar reasoning of specialization and concentration. He indicated four basic influences of clustering as the presence of related industries and associated institutions, the sophisticated and specialized demand, qualified and specialized factors, and local contexts for firm strategy and rivalry. Lucas (1993) also underscored

[34] These three factors are similar with the Marshall's model. Krugman (1991b) also mentioned the difficulty of measuring and tracking knowledge spillovers, mentioning that "knowledge flows are invisible, they leave no paper trail by which they may by measured and tracked" (Krugman, 1991b, p.53).

increasing returns from agglomeration as an only convincing reason for the existence of city, where communication is facilitated.

Meanwhile, there is other counterpart explanations which tell us the concentration may lead to congestion, pollution, high cost for transportation or land use (see, for example, Abdel-Rahman and Fujita,1990; Helsley and Strange, 1990; Ottaviano and Puga, 1998; Duranton and Puga, 2000).

Agglomeration economies, which are external economies as generating additional benefits through sharing resources by being located closely, were broken into "localization" and "urbanization" economies, which originally developed by Hoover (1961). Localization economies mean externalities generated within same industry as the specific advantages based on the specific demands within an industry, and therefore is related to specialization. Localization economies emphasize the sharable characteristics within same industry, such as similar technology, similar inputs, similar customers, and similar workers. Sullivan (1993) explained that localization economies occur when the production costs in a particular industry decrease according to the increase of its total output, due to three key reasons which are scale economies in intermediate inputs, labor-market economies, and communication economies.

Urbanization economies mean externalities generated from different industries and therefore are related to diversity. Urbanization economies occur when the production costs of an individual firm decease according to the increase in total output of the urban area based on similar reasons with localization economies but in the broader terms, for example, the provision of overall business services and public services, and a stable level of total employment (Sullivan, 1993). He explained that business services include banking, insurance, real estate, hotels, building maintenance, printing and transportation and public services include highways, mass transit, schools, and fire protection. Urbanization economies highlight complementary relationships of different economic activities and the necessity of their co-existence and therefore, an overall size of region is important for this type agglomeration.

More specifically, the debates of diversity and specialization are made focusing on the role of knowledge spillovers. Following Marshall (1890), Arrow (1962a), and Romer (1990), the possibility of imitation and improvement through knowledge spillovers, has been paid attention as the major benefit of specialization, called as the MAR-type specialization. Porter (1990) also supported concentration of similar industries for economic growth and explained that the growth is promoted through competitions between the firms within industry because these kinds of competitions reinforce knowledge spillovers by using an example of Italian ceramic and gold jewelry industries. Contrarily, Jacobs (1969) emphasized the mix of different ideas and different sources of technology. She argued that various possibilities to encounter diverse knowledge and technology are necessary for increasing and opening the opportunity of innovation. Consequently, she insisted on the importance of interactions (including competitions) across industries not just within industry. Her well-known example is the starting of the bra industry in New York city.

Empirically, Glaeser, Kallal, Scheinkman and Shleifer (1992) provided evidence of Jacob-type agglomeration effects based on their empirical finding which showed us that the faster growth of industries' employment are observed in the less-specialized, small firm-dominant cities rather than the overrepresented cities by analyzing six largest industries in 270 cities between 1956 and 1987. Harrison, Kelley and Gant (1996) also found the importance of diversity, or urbanization economies for adopting new technology in their study of the localization and urbanization effects on innovative activities of metalworking industry.

Contrarily, Henderson (2003) provided an evidence of MAR-type specialization but not Jacob-type diversity by estimating an effect of two types of scale externalities (measured by the count of own industry plant and the number of employment in other industries respectively) on productivity growth for the high-technology industry. Meanwhile, Henderson, Kuncoro and Turner (1995) found positive effects of knowledge spillovers both within and across industries on city specialization by using employment growth data and also indicated these effects depend on the product cycle. In detail, at the initial stage of product development they found interactions across industries (i.e. Jacobs-type agglomeration) are more frequent, but during maturation period interactions within industry (i.e. MAR-type specialization) are observed more prevalently in order to generate more significant effects.

More broadly, Dumais, Ellison, and Glaeser (2002) provided an empirical finding of agglomeration effects on employment from various factors, in particular, a strong evidence of labor pooling, from the metropolitan area data between 1972 and 1997. Rosenthal and Strange (2001) also provided overall assessments of agglomeration, which include effects of knowledge spillovers, labor market pooling, input sharing, product shipping costs, and natural advantage at the state, county and zip code levels from U.S. manufacturing industries at the year of 2000. In both of those researches, the geographical concentration is measured by the Ellison-Glaeser index mentioned below. Ciccone and Hall (1996) found employment density (measured at the county level) promotes the growth of labor productivity (measured at the state level), implying that density is a major force for increasing returns through agglomeration effects.

For measuring the degree of specialization of a region, several indices, such as the Gini coefficients, Herfindahl, the Ogive and the entropy, are widely used. These indices are calculated based on the number of industries or the output of each industry.[35] Henderson, Kuncoro and Turner (1995) used the Herfindahl index. In particular, Ellison and Glaeser (1997) developed the modified index for spatial specialization based on the Herfindahl index. The Ellison-Glaeser index captures the geographical concentration of employment, linked with the distribution of establishments and this index is popularly adopted in many agglomeration literatures including Dumais, Ellison and Glaeser (2002) and Rosenthal and Strange (2001). In addition, establishment density, employment density, or population density, as a simplified one, could also be an alternative measurement of agglomeration (see, for example, Ciccone and Hall, 1996). Because the degree of specialization can be altered according to how to define a level of industry or region, interpretations should be carefully placed. Table 3 summarizes the alternative measurements of the degree of regional specialization and their applications.

In sum, agglomeration economies have been recognized as one of important factors for regional economic growth through generating external economies as additional benefits through sharing resources by being located closely. This kind of benefits from external economies can be divided by urbanization economies generated from interactions across industries and localization economies, generated from interactions within industry. Indeed, these two types of agglomeration economies are supported by different roles of knowledge spillovers as well. Which kind of agglomeration economies more generates regional economic growth is still in debates based on mixed empirical findings mentioned above.

[35] For the extended surveys of measurements, refer to Siegel, Johnson, and Alwang (1995), Wagner (2001), and Wundt (1991).

Table 3. The alternative measurements of regional specialization

Possible measurements	Definition	Application
Gini coefficients (Audretsch and Feldman, 1996)	$\frac{\text{The value-added in an industry and state}}{\text{The national value-added for the industry}}$ (normalized by state share of total manufacturing value-added)	- Krugman (1991a) - Audretsch and Feldman (1996)
Herfindahl Index	$H = \sum_{j=1}^{N} Z_j^2$ where j means firms in an industry, Z_j means the employment share of the jth firm.	- Henderson, Kuncoro And Turner (1995)
Ellison-Glaeser Index (Ellison and Glaeser, 1997)	$\gamma = \frac{G-(1-\sum_i x_i^2)H}{(1-\sum_i x_i^2)(1-H)} = \frac{\sum_{i=1}^{M}(s_i - x_i)^2 - (1-\sum_i x_i^2)H}{(1-\sum_i x_i^2)(1-H)}$ where G means spatial concentration, H means Herfindahl index, M means geographic areas, s_i means industry employment in each area, x_i means total employment in each area.	- Ellison and Glaeser (1997) - Dumais, Ellison, and Glaeser (2002) - Rosenthal and Strange (2001)
Share at the city level relative to the share at the national level	$\frac{\text{Industry employment in city /total employment in city}}{\text{Industry employment in US /total employment in US}}$	- Glaeser, Kallal, Scheinkman, and Shleifer (1992)
Employment density	Number of employee per area	- Ciccone and Hall (1996)
Establishment density	Number of establishment per population	

3.2.5.3 Competition

Third, competition is also considered as being essential for regional economic growth. It is believed that stimulating and contesting each other could accelerate improvement of economic outcomes. Porter (2000) stressed the role of competition and rivalry for regional economic growth within clusters, defined as "geographic concentrations of interconnected companies, specialized suppliers, service providers, firms in related industries, and associated institutions (e.g., universities, standards agencies, trade associations) in a particular field that compete but also cooperate" (Porter, 2000, p.15). He argued that competition and rivalry encourage sustained investing, innovating, improving productivity, and upgrading. Competition is also essential for “technological progress” and “creative destruction” within Schumpeter-type innovation and economic growth.

The importance of competition could be understood through the relationship with knowledge spillovers. Namely, Porter (1990) favored competition for augmenting innovative activities through knowledge spillovers that increase economic growth as a result. Emphasizing diversity, Jacobs (1969) believed competition between diverse actors encouraged innovation. Contrarily, Romer (1986) and Arrow (1962a) emphasized the negative aspect of competition. For example, knowledge spillover through competition could be shared without appropriate compensation due to the property of one of public goods.[36] Therefore, monopoly is good for growth because it raises incentives for innovation by internalizing possible externalities. In this sense, the argument of competition and monopoly is closely related to the argument of diversity and specialization. Chinitz (1961) also underscored that the competition is essential for economic growth by encouraging learning, innovation, and entrepreneurship among diverse and small-sized firms. In summary, competition can produce opposite outcomes with respect to knowledge spillovers as follows: competition could enrich knowledge spillovers and encourage innovation based on possible gains to imitators, ultimately resulting in productivity growth and economic growth, while competition could provide incentives for under-investing in R&D and discouraging efforts for innovation due to possible losses to original innovators, consequently undermining productivity growth and economic growth.

The degree of competition could be measured by the number of firms or the number of firms divided by the size of the industry as a simple one (Glaeser, 2000). Glaeser, Kallal,

[36] There are some efforts to internalize the externalities of knowledge, such as a patent system and intellectual rights. However, it is believed that knowledge spillover still exists with these policies in force.

Scheinkman and Shleifer (1992) measured competition by the number of firms per worker within an industry and a city relative to the number of firms per worker within the industry in US, and found the positive relationship between competition and economic growth. Feldman and Audretsch (1999) used the same index as Glaeser, Kallal, Scheinkman as and Shleifer (1992) for measuring competition,[37] and found that competition and diversity promote innovative activities.

In sum, the above empirical studies found that competition and diversity rather than monopoly and specialization promote regional economic growth by examining the role of knowledge spillovers. Because the arguments of competition are closely related to the arguments of specialization and diversity, the selected measurements, findings, and following implications are quite relevant as well.[38]

3.2.5.4 Government policy

Whether government policy affects regional economic growth is also controversial. While there are some strong beliefs of market distortions resulted from national level government policies (Glaeser, 2000), there are some variations in analyses of the effectiveness of local governments' policies. Under the situation of increasing popular utilization of regional economic development strategies to cope with the severe inter-regional competitions, the necessity of understanding their consequences is growing. However, fundamental difficulties of

[37] Feldman and Audretsch (1999) also used the same index as Glaeser, Kallal, Scheinkman as and Shleifer (1992) for measuring specialization, which is the share of total employment by an industry employment in a city relative to the share of total employment by the industry employment in the US. In addition, they added the index for measuring diversity by modifying the above specialization index, which is the share of total employment by science-based industries employment in a city relative to the share of total employment by science-based industries employment in the US.

[38] The importance of competition for economic growth also could be considered at the firm level. For example, Feldman and Audretsch (1999) found the importance of diversity rather than specialization at the firm level by using the non-linear economic model developed by Scherer (1983).

evaluating the effectiveness of regional development strategies are also well-known. These difficulties include a difficulty of separating effect by each strategy, a difficulty of separating effect from other relevant factors, and defining a group of recipients, an impact area, or a proper time period. Here, what we need is the more appropriate evaluation for improving and modifying policy tools instead of just abandoning the entire program with uncertainty.

Brown, Kathy and Taylor (2002) categorized analyses of regional economic development policies in three ways: 1) finding effects on regional economic growth and firm location,[39] (2) findings relationships between local policy and regional wage differentials,[40] and 3) finding the degree of public capital supplies through relationship between regional factor inputs and factor outputs.[41] Then, by integrating these three threads, they found that state and local policies have more affected to the private capital-to-labor ratio rather than the private economic output and overall increase of government discourages economic growth in the private sector.

Besides the studies mentioned above, the following research showed some evidence of policy impacts despite the difficulties mentioned above, and the findings were somewhat mixed. As overall assessments, Glaeser, Scheinkman and Shleifer (1995) and Rauch (1995) found little evidences of positive regional policy effects, measured by the increase in income and population, and the increase in the share of infrastructure investment respectively, while Rappaport (1999) and Vidgor (1998) showed strong evidences of negative regional policy effects, measured by the decrease in population and the decrease in tax base respectively.[42]

---

[39] See, for example, Papke (1991) and Gray (1997).

[40] See, for example, Beeson and Eberts (1989) and Haughwout (2002).

[41] See, for example, Garcia-Milà, McGuire and Porter (1996) and Ferland (1999).

[42] Glaeser (2000) explained that these negative effects are linked with low level human capital accumulation resulted from regional policies.

More specifically, the effectiveness of tax benefits, in terms of the attempts to induce private firms through a set of economic incentives among the competitive regions, is still inconclusive as well. The fundamental question about these local level economic development policy tools is whether they generate a positive-sum game as creating an additional economic outcome *or* a zero-sum game as a simple relocation of individual firms from one region to the other region. Garcia-Milà and McGuire (2001) argued that the tax incentives work efficiently by generating positive externalities as a form of agglomeration economies if they are offered as a tax rate below the benefit tax level. There are some empirical findings related to this issue.[43] Holmes (1998) provided some positive effects of right-to-work laws on attracting industry. Saiz (2001) found the entrepreneurial economic development strategy, which is a combined measure of various state policy tools created by author, has a positive growth effect on manufacturing employment but, no effect on state total gross product and unemployment. Bartik (1991) and Wasylenko (1997) found the little evidence of effects on employment of local taxes. However, Mark, McGuire and Papke (2000) found the statistically significant and negative effects of personal property and sales taxes on employment growth in Washington D.C. metropolitan area during 1969 and 1994.

In sum, the effects of regional government policy on regional economic growth are somewhat ambiguous depending on the definition of policy programs and the presumed effects. Fundamentally, the evaluations of regional economic development programs encounter some difficulties of separating effect by each strategy, separating effect from other relevant factors, and defining a group of recipients, an impact area, or a proper time period. The sophisticated

[43] Empirical evidences of R&D-relevant policies including state-level and federal-level tax credits are separately examined at chapter 4, section 4.4.

methods and the appropriate research designs for dealing with these difficulties are necessary for understanding the effectiveness of regional level government policies and providing the future suggestions of the better policy utilizations.

3.2.5.5 Initial economic condition

Initial economic condition could affect future economic growth in several ways. From the force of the convergence, the high initial condition would be lowered and vice versa, which subsequently results in reducing inequality. Meanwhile, there is a possibility of the divergence, which implies a region having the higher level of economic situation such as the higher income and the larger population, tend to keep gaining and growing faster, while a region having the lower level of economic condition tend to keep gaining and growing slower, ultimately resulting in the deeper disparity. Profit cycle theory, developed by Markusen (1985), also provided the possible linkages of initial economic conditions, such as various economic factors and market demands, with future economic growth of employment, output, investment, occupational composition, and geographical concentration. Chinitz (1961) also provided the possible effects of initial economic conditions by contrasting different industrial structures and the following different economic growth of New York and Pittsburgh.

In addition, some argued that city growth rates could be determined by initial city sizes based on Gibrat's law (Gibrat, 1931). Zipf's law tells us that city sizes follow a rule of pattern by their rank, which is, in detail, the *n*th ranked city has 1/*n* the population of the largest city (Zipf, 1949). These two laws are quite relevant (see, for example, Gabaix, 1999a and 1999b). Following Zipf's law, the second largest city has one-half the population of the largest city and the third largest city has one-third that population. Within a certain time period, steady change instead of huge jump is mostly expected, and therefore the previous condition could provide some useful clues about expecting the future and detecting possible shocks from outside if any exists. Moreover, if we observe continuous changes over time, we can estimate possible gains in the future.

There are some mixed empirical findings about relationship between initial economic conditions and the following growth of them. About income, Barro and Sala-I-Martin (1991) provided strong evidence of convergence, which means the negative effect of initial high income on the following income growth by using the state-level data in the US as well as the other country-level data. However, Glaeser, Scheinkman and Shleifer (1995) found no clear evidence of income convergence by using the city-level data between 1960 and 1990. About employment and population, there are no clear evidences of convergence (see, for example, Eaton and Eckstein, 1997; Glaeser, Scheinkman and Shleifer, 1995), while there are some strong evidences of divergence, namely the persistent population growth with the high level of initial population (see, for example, Blanchard and Katz, 1992; Glaeser, Scheinkman and Shleifer,1995; Beeson, Dejong and Troesken, 2001).

Indeed, Nitsch (2005) tested the Zipf's law by doing meta-analysis of twenty-nine previous empirical studies and concluded that the distribution of city size does not exactly follow the Zipf's law and rather the estimates tend to be large, implying that the population of the second largest cities tends to be greater than one-half the population of the first largest city, and the population of the third largest cities tends to be greater than one-third the population of the first largest city and so on.

In sum, the possible effects of initial economic conditions on future economic growth are recognized by either the divergence or the convergence. Both of these possibilities are found in empirical analyses depending on the measurement of economic growth.

In this section 3.2.5, I examined the five important factors to affect regional economic growth. These are human capital, specialization/diversity, competition, government policy, and initial economic condition. Understanding the mechanism of their influences in particular at the regional level and recognizing possible measurements, provide the fundamental basis of developing evaluation strategies, in particular, developing plausible other explanations and the following empirical specifications. The relevance of these factors with possible economic outcomes from the state R&D tax credits will be examined in Chapter 6.

In this chapter, I explored the theoretical literatures relavant to R&D. They are broadly categorized by R&D-based economic growth theories and regional economic development theories. The first set of theories tell us why R&D is important, to what extent R&D is performed and linked with economic growth, and why government intervenes private R&D and so on. These theories provide the basis of the legitimization of state R&D tax credits and the explanations of their appropriateness, effectiveness, and possible consequences. The second set of theories show us to what extent private R&D plays different roles in promoting regional economic development within several different perspectives. Then, I explored the important factors for regional economic development especially with the extensive analyses of empirical findings. These factors are useful for constructing rival hypotheses and selecting either covariates or matching variables within empirical specifications of this study.

In the next chapter, I discuss the historical utilization of R&D tax credits in the US at the federal level and the state level. I also discuss the previous evaluations of these policies as well.

4.0 R&D TAX CREDITS AND POLICIES

In this chapter, I examine state R&D tax credits in terms of their application rules, their uses in the U.S., and previous analyses of their effectiveness. Because most state R&D tax credits are some modifications of the federal R&D tax credit, the development of the federal R&D tax credit is also briefly introduced. Then, the development of the state R&D tax credits is discussed focusing on their different utilizations by state. Finally, previous evaluations are examined and those evaluations include evaluations for state R&D tax credits, the federal R&D tax credit, and direct government funding.

The R&D tax credit is one government policy used for promoting economic growth by encouraging private R&D activities. Indeed, as an indirect government intervention, state R&D tax credits are also one of the popularly utilized regional economic development strategies. Based on these characteristics, their basic rules and possible effects are investigated. Then, the overall features of indirect government intervention are assessed with the comparison of direct government intervention. From different roles of these two kinds of interventions, complementary utilizations of them are suggested.

In practice, the R&D tax credit was introduced in the United States at the federal level by the Economic Recovery Tax Act (ERTA) of 1981. Since then, this tax credit law also has been enacted at the state level and, by 2003, 38 states adopted their own R&D tax credit programs in somewhat different ways in terms of the base of credits, the operation year, the geographical distribution, the application rule, and the appointment method of state corporate tax. Because most state-level R&D tax credit programs were modeled after the federal law, the federal R&D tax credit is briefly examined first. The federal R&D tax credit has only been temporary. It was renewed and modified several times.

Some selected states' laws, which are California, Massachusetts, Michigan, Texas, and Washington, are investigated in detail. These comparisons provide fundamental bases on the following empirical analysis of the effectiveness of this program, in that the program is evaluated through comparing states having tax credits and states not having them, based on quasi-experimental methods.

Before going further, previous studies of analyzing the effectiveness of state R&D tax credits are examined within this chapter. There are few previous evaluations of state level R&D tax credits. They did not provide quantitative analyses and instead used descriptive and anecdotal approaches. However, within these analyses, the important issues related to the tax credits are pointed out. Indeed, these evaluations were performed for each state law, and therefore they did not yield comprehensive analyses. The selected evaluations for California, Virginia and Washington, are explored in detail. With this scarcity of evaluating state R&D tax credits, previous evaluations of R&D relevant policies, in particular the federal level R&D tax credit and direct R&D funding, are explored. Compared with evaluations of state tax credits, evaluations of federal tax credits and direct R&D funding are more numerous and provide some useful references for figuring out important relevant factors, available datasets, and possible specifications for quantitative analyses.

4.1 OVERVIEW

First, I investigate the basic application rules of the R&D tax credit. Although there are some variations for operating R&D tax credits by state, the application rules mentioned in this section are popularly applied ones across states modeled after the federal R&D tax credit. The R&D tax credit intends to increase private R&D spending by allowing firms to deduct their R&D expenditures from corporate taxes. By definition, applicable R&D expenditures are qualified research expenses and basic research payments. Qualified research expenses are defined as expenses "in the experimental or laboratory sense if they are for activities intended to discover information that would eliminate uncertainty concerning the development or improvement of a product" based on the section 174 of the federal tax code (Treasury regulations, sec. 1.174-2(a), requoted by OTA, 1995a, p. 13). Qualified research expenses consist of in-house research expenses (including wages and supplies) and contract research expenses, which exceed the base amount as an incremental way. The detailed definition of each item is summarized in Table 4.

Table 4. The definition of eligible R&D expenditures

Qualified research expenses	In-house research expenses; Wages: amount paid to employees for qualified services Supplies: tangible property other than land or depreciable property
	Contract research expenses; 65 percent of amount paid or incurred by the taxpayer to any person (other than an employee of the taxpayer) for qualified research.
Basic research payments	Amount paid in cash by a corporation to any qualified organization for basic research

Source: CCH Tax Research Network and Rashkin (2003)

Rashkin (2003) categorized qualified research expenses in detail as follows: 1) compensation of researchers, 2) overhead, 3) contract research, 4) supplies, 5) capital expenditures, 6) license of intangible assets, and 7) software. He also pointed out that eligible research expenses for claiming the credit are limited to research expenses relevant to direct innovation and therefore research expenses relevant to product development are excluded. The interpretation of the eligible research expenses are still in debate. Most states followed the federal level definition of qualified R&D expenses with limiting the research expenses within a state. The modifications include the different tax credit rate and limitation of the research expenses based on the industry and area.

By allowing qualified R&D expenditures that only exceed a certain base level, this program may possibly avoid a windfall effect which means providing credits for the activities that would have taken place in the absence of the credit. Based on the federal law of Code Sec. 41, which is mostly followed at the state level,[44] twenty percent of applicable R&D expenditures can be applied to the R&D credit. It shows us that all of the R&D expenditures are not eligible, and therefore the R&D tax credit is only a partial deduction of R&D expenditures.

From the above application rules, state R&D tax credits affect increases in employment and payroll by deduction of in-house research expenses. They also encourage increases in R&D activity and employment of multiple firms at the same time by deduction of contract research expenses.

Creating jobs is a primary objective of overall regional economic development strategies.[45] Making job growth as the priority also can be understood in the context of regional competition for economic gains. Rashkin (2003) pointed out employment opportunities related to R&D are mostly high-wage and professional, and they also can be linked with creating other

[44] Some states have made some modifications, for example, the changes in percentage of eligible amounts.

[45] Some states (i.e. Washington and Michigan) explicitly demonstrate in the law that the purpose of R&D tax credit is to increase employment.

manufacturing jobs. Together with these effects, it is also expected that overall outcomes of firms such as innovative activity and productivity will be improved by increasing R&D, and thus, this is another purpose of the R&D credit by supporting basic research.

Meanwhile, as a regional economic development policy, tax incentives might affect a firm's location decisions. However, it is broadly believed tax incentives have had only marginal effects for these decisions (for example, Greenbaum and Engberg, 1998; Papke, 1993). It is also true in the sense that most states provide some kinds of tax incentive packages, and these are becoming more similar rather than more diverse (Elling and Sheldon, 1991).

As an alternative of R&D tax credits which are related to mostly corporate income tax, there is a sales and use R&D exemption for purchases of inputs to R&D. Because a sales and use R&D exemption is mostly applied to only purchasing machinery and equipment directly related to R&D, Hall and Wosinska (1999) advantaged tax credits due to a possibility of broader applications. For example, through tax credits firms can deduct every eligible R&D expenditures, including personnel costs and spending for basic research, which are important for R&D activities, from their corporate taxes. However, they also pointed out that start-up companies could favor a sales and use R&D exemption because they could not make taxable earnings for receiving tax credits in their initial stages and it is hard for them to achieve a quick depreciation of equipment expenditures, which is necessary for obtaining benefits in case of an incremental R&D tax credit that is most popular among state laws.

The legitimacy of the R&D tax credit can be found from overall rationale for tax credits. A tax credit, as one of the tools performed by government for encouraging private investment, is generally suggested for achieving an optimal point of social benefits, which otherwise cannot be reached only by private investment. It comes from a market failure, which means that a private market can't work properly for accomplishing an optimal point for social benefits.[46]

Based on the above rationale, government provides grants and subsidies to a private industry or establishes public-funded research centers as direct intervention, or provides tax incentives as indirect intervention, for encouraging private R&D activities and ultimately promoting R&D growth. These two different kinds of policy tools could be distinguished by decreasing the marginal cost of R&D investment through tax incentives and increasing marginal rate of return of R&D through direct subsidies (David, Hall and Toole, 2000).[47] In addition, indirect intervention is more related to reduce the cost while direct intervention is more related to provide the less-profitable researches which tend to be undersupplied within a private market mechanism.

As Bozeman & Link (1984) illustrated, there are some advantages and disadvantages of each policy intervention. At the position of indirect intervention, tax incentives allow relatively more autonomy to private decision-makers as well as require less administration. Indeed, tax incentives intend to provide benefits based on economic outputs from the private sector's own decision and its own market-oriented demand, while direct subsidies might provide benefits created and planned by government, not directly determined by the private sector. In terms of operational flexibility, tax incentives can be long-lasting with no necessity of annual budget review and they can sustain from various political positions. Besides the psychological

[46] The reason that is involved in a market failure can be classified as following. One is from monopoly. For establishing an appropriate market system, competition is required for determining the optimal point of price and amount of products. Second comes from public goods, which usually create benefits for multiple sectors, but there is no incentive for providing them. Accordingly the provision of public goods is less attractive for private sector. Third one is information failure, which means participants of a market can't capture necessary information in terms of amount and timing, and consequently, it is difficult to make a decision. Others are incomplete markets, in other words, externalities (Stiglitz, 1988).

[47] David, Hall and Toole (2000) also explained different effects between government R&D contracts and grants.

advantage for achieving tax incentives is another possible benefit. Meanwhile, the possibility of unintended windfalls and undesirable inequities is pointed up as one of the disadvantages of tax incentives. Private firms might raise their R&D spending without tax credits, which means over-subsidy. Indeed, some firms, which can not be qualified in the specific rules of tax incentive programs, might be excluded. The problem is that big firms usually are beneficiaries of the programs and new small firms tend to be excluded with high possibilities (see, for example, Sigalla and Viard, 1999). Tax credits also affect government budget controls and undermine accountabilities. There is other possibility of the stricter utilization of tax incentives by specifying a lot qualifications and the more flexible utilization of grant programs by applying generous rules.

On the contrary to tax credit as indirect interventions, direct government interventions could affect private R&D activities in two different ways, either "stimulating" or "crowding out" (David, Hall and Toole, 2000). Stimulating private R&D means public R&D funding could generate the increase of private R&D funding as a complementary relationship and ultimately result in the productivity growth. Stimulating private R&D could be expected if public R&D successfully perform the specific R&D area that is essential for future productivity growth but might be avoided by the private sector based on high risk and uncertainty. Any applicable findings from public R&D could encourage private R&D for better utilization or practical development of them. Tax credits are expected to generate stimulating private R&D by giving an incentive to increase private R&D for reducing R&D costs.

Meanwhile, crowding out private R&D means public R&D investment simply substitutes for private R&D investment in the sense that without public R&D investments, private firms might make these investments at their own expense and therefore, with public R&D investments, private firms might decrease their own R&D investments. If public R&D activities are involved in R&D projects which the private sector also intend to do by its own, the private sector could simply utilize the findings without its own efforts. However, in the case of tax credits, they are more related to reducing R&D costs based on increased taxable incomes and R&D spending and therefore, crowding out effect is generally not expected.

These possible opposite effects depend on the characteristics of R&D projects chosen. For example, R&D projects, which are long-term based and relevant to basic research and therefore benefits from which are unclear, might be avoided from the private sector, while R&D projects, which are short-term based and relevant to applied research and therefore benefits from which are expected as an immediate one might be easily performed by the private sector.

More recently Klette, Møen and Griliches (2000) pointed out the limitation of R&D tax credits due to low response elasticity. However, Hall and van Reenen (2000) favored R&D tax credits but un-favored direct government intervention because there is a high uncertainty of future demand on knowledge in spite of the better opportunity for doing targeted projects for higher social return and we already have much experienced government failures. David, Hall and Toole (2000) also indicated R&D tax credits do not result in a crowding out effect which is mostly expected as a negative result of direct subsidies. However, they also mentioned the possible tendency of inclining short-term R&D investments for reducing taxes from immediate increased earnings and this tendency could lessen long-term investments which expectedly result in much larger spillover effects, and consequently much higher social rates of return. This tendency might be less shown in direct intervention, even though there is a possibility of reliance on short-term projects for obtaining instant and visible effects of policy implementation.

Table 5 summarizes the possible advantages and disadvantages of indirect intervention and direct intervention to support private R&D. As Mamuneas and Nadiri (1996) argued, these two different policies have complementary relationships and therefore, the balanced utilization is important for satisfactory policy effects. Namely, tax credits encourages the increase of private R&D investment through reducing the cost, while direct intervention complements to private R&D activities through providing the less profitable and long-term based R&D activities that tend to be avoided by the private sector.

Table 5. Comparison of indirect intervention and direct intervention to support private R&D

	Possible advantage	Possible disadvantage
R&D tax credit as an indirect intervention	- Stimulating effect - No crowding out effect - Autonomy - Operational flexibility	- Possibility of inclining short-term projects - Unintended windfall: over-subsidy - Undesirable inequality: excludability of small firms - Low response elasticity
Subsidy as a direct intervention	- Stimulating effect - Less-profitable project - Long-term projects - Complementing to private R&D	- Crowding out effect - Possibility of performing less-efficient project due to government failure

Based on these theoretical rationales for R&D tax credits, I examine the actual utilizations for R&D tax credits in the U.S. in the next section. Because most state R&D tax credits are some modifications of the federal R&D tax credit, the federal law is examined first in the next section and then, the discussion of state R&D tax credit programs follows.

4.2 THE FEDERAL LEVEL R&D TAX CREDIT

In the US, the R&D tax credit begun to be implemented at the federal level and the federal R&D tax credit has been placed in an important model for developing state R&D tax credits. As one type of indirect intervention, the R&D tax credit in the federal level was introduced at the Economic Recovery Tax Act (ERTA) of 1981 with the purpose of inspiring additional private R&D activities.[48] This law was primarily enacted as a temporary law, and thereafter it has been renewed and extended eleven times. The temporary utilization is aimed at examining the appropriateness of this credit based on the uncertainty of its effect as well as the legitimacy of policy intervention to private R&D activity, and ultimately finding more efficient application rules. Usually the credit was extended before expiration, and in the case that it was renewed after expiration, it applied retrospectively. One exception was the year of 1995, namely the Tax Reform Act of 1996, which was not applied to 1995. Recently, the federal R&D tax credit was extended for 18 months by the Working Families Tax Relief Act of 2004 and this law was expired at the end of the year of 2005. Because the federal R&D tax credit is still temporary, this law still remains as one of controversial issues in particular at the time close to the expiration.

Basically the R&D tax credit is applied in an incremental way, which means it is applied to only the amount in excess of the amount of base year. Among eleven renewed legislations, the principle didn't change much but there were two major changes in 1986 and 1990. By the Tax Reform Act of 1986, the tax credit was applied to only "qualified Research expenses," that usually include the basic research and the applied research depending on interpretation.[49] By the

[48] At the first time, this law called as the R&E credit (Research and Experimental credit) instead of the R&D credit (Research and Development credit).

[49] Based on Internal Revenue Code and Treasury regulations, Sigalla and Viard (1999, p.3) defined the qualified research

Omnibus Budget Reconciliation Act of 1990, the base year changed from previous years to fixed years as 1984-1988 and for start-ups the fixed rate of 3% was applied. This change is based on the critique which tells us that with the base year as previous years, increasing the current year spending makes it difficult to earn a credit in the following year. However, the rule of fixed years is also problematic because each firm has a different experience of those years. Namely, some firms with high level of R&D investment between 1984 and 1988 relative to current R&D investment, are expected to be highly denied the credits, and vice versa. In 1996, the Alternative Incremental Credit (AIC) was added as an alternative method of calculating the base amount and this alternative method provides the benefit to the firms experiencing a rapid sales growth. These firms are mostly categorized by small high-technology industries, more specifically pharmaceuticals, computer manufacturing and electrical and electronic machinery manufacturing (Hall and Wosinska, 1999). Table 6 illustrates the detailed features of changes in federal level R&D tax credit policy.

---

expenses as follows;

Research must consist of a "process of experimentation" in engineering, physics, biology or computer science and must seek "technological" information not commonly known to skilled professionals. The research effort need not be successful. The information sought must be useful in developing a "new or improved" business product or technique and must relate to function, performance, reliability or quality, and not style. The credit does not apply to "reverse engineering," market research, routine quality control or research following commercial production.

These criteria is somewhat subjective and therefore debates on the interpretation are involved in. Generally development is excluded due to the expectation of little spillover effect. (Sigalla and Viard, 1999)

Table 6. The major changes of the federal level R&D tax credit policy

Period	Legislation tax act	Credit rate	Corporate tax rate	Definition of base	Qualified expenditures
July 1981 to Dec 1985	The Economy Recovery Tax Act of 1981	25%	46%	Maximum of previous 3 year average or 50% of current year	Excluded: R&D done outside US and funded by others, R&D in Humanities/Social science
Jan 1986 to Dec 1988	The Tax Reform Act of 1986	20%	34%	Same	Narrowed definition to "technological" R&D, excluded: leasing
Jan 1989 to Dec 1989	The Technical and Miscellaneous Revenue Act of 1988	20%	34%	Same	Same
Jan 1990 to Dec 1990	The Omnibus Budget Reconciliation Act of 1989	20%	34%	1984-88 R&D to sales ratio times current sales (max. ratio of 16%), or 3% for start-ups	Same
Jan 1991 to Dec 1991	The Omnibus Budget Reconciliation Act of 1990	20%	34%	Same	Same
Jan 1992 to June 1992	The Tax Extension Act of 1991	20%	34%	Some modifications for start-ups	Same
July 1992 to Dec 1993	The Omnibus Budget Reconciliation Act of 1993	20%	34%	Same	Same
Jan 1994 to June 1995		〃	35%	〃	〃
July 1995 to June 1996	N/A	0%	35%	N/A	N/A
July 1996 to June 1997	The Small Business Jobs Protection Act of 1996	20%	35%	1984-88 R&D to sales ratio times current sales (max. ratio of 16%), or AIC*	Same
July 1997 to June 1998	The Taxpayer Relief Act of 1997	20%	35%	Same	Same
July 1998 to June 1999	The Omnibus Consolidated and Emergency Supplemental Appropriations of 1998	20%	35%	Same	Same
July 1999 to June 2004	The Tax Relief Extension Act of 1999	20%	35%	Same	Same
June 2004 to Dec 2005	The Working Families Tax Relief Act of 2004	20%	35%	Same	Same

Notes: *AIC - Alternative Incremental Credit
Source: Hall (1993), OTA (1995a), Hall and Wosinska (1999), and updated by author.

Along with the R&D tax credit, the federal government has intervened in private R&D directly through federal R&D funding since the World War II, which implies it has a long history for supporting R&D. As one way of intervention, the federal government has made financial support to industry, universities/colleges, and nonprofit organizations. There was a significant shift of resource allocations, which means that the federal government support for R&D was focused on the defense-related industry for 40 years and then it was redirected to the financial, human, capital resources in the 1990s (NSF, National Science Board, 1993).[50] Although the federal government's share of total R&D funding in the US has decreased over time, namely 66.8 percent in 1964, 50 percent in 1979, and 25.1 percent in 2000 (NSF, Division of Science Resource Studies, 2005), the role of the government-performed R&D is still significant. Namely, its R&D activity plays an important role because public R&D activity can stimulate and complement private R&D activity. They also are closely interrelated through knowledge transfer and knowledge spillover.

As the other way of intervention, the federal government has performed R&D by itself, establishing the "Federally Funded Research and Development Centers" (FFRDCs). The FFRDCs are defined as the "R&D-performing organizations that are exclusively or substantially financed by the federal government and are supported by the federal government either to meet a particular R&D objective or, in some instances, to provide major facilities at universities for research and associated training purposes" (NSF, Division of Science Resource Studies, 2004, p.9). They were called the "Federal Contract Research Centers" until 1967. Some examples are Lincoln Laboratory Logistics Management Institute by Massachusetts Institute of Technology,

---

[50] Between 1970 and 1997, the Department of Health and Human Services increased its share of federal funding by 17 percent however the Department of Defense decreased its share by 14 percent (NSF, Division of Science Resource Studies, 1999).

National Defense Research Institute by RAND Corporation, and Lawrence Berkeley Laboratory by University of California.[51] Twenty three FFRDCs were established in 1950 and thereafter the number of the FFRDCs increased by over 60 in early 1960s and peaked at 74 in 1968 (OTA, 1995b). Then, the number of centers has been decreasing by approximately 30 and these corresponded mainly to the defense-related FFRDCs (OTA, 1995b). In detail, the number of the defense-related FFRDCs substantially changed from forty to ten, while the number of other FFRDCs remained at approximately 20 for 45 years (OTA, 1995b).

Within the extensive government intervention to encourage private R&D at the federal level discussed above, the federal R&D tax credit has been an important policy, even though it is still temporary law and the debates regarding to its effectiveness still remains. In next section, I discuss the utilizations of state R&D tax credits. Their utilizations are different across states and their effectiveness is also still in question. Therefore, understanding these differences can provide a fundamental basis of evaluating their effectiveness.

4.3 STATE LEVEL R&D TAX CREDITS

As mentioned above, the state-level R&D tax credit policies have taken a different form by the state level. For understanding these differences of state's tax credit laws, the next four criteria are chosen: (1) the base for the credit, (2) the operation year, (3) the application range, and (4) the state appointment method for corporate tax. Then, some selected state laws (i.e. California, Texas, Massachusetts, Michigan and Washington) are examined in detail. These states are the

[51] For the detailed list of current and historical FFRDCs, see OTA, 1995.

states performing a high level of private R&D in terms of the amount of R&D spending and their tax credit laws has been developed in a similar way.

4.3.1 The basic features of the state policies

In this section, I examine the following features for comparing each state's tax credit law. First, state R&D tax credits are distinguished by the base for the credit. Most states use R&D expense for calculating credits while some states use employment. Second, these tax credit programs were enacted in different time period, from 1980s to 2000s. Third, most tax credits are utilized generally to be eligible for every qualified R&D expenditures within states, while some are utilized narrowly to be eligible for R&D expenditures for limited geographical areas or limited industries. Fourth, each state has the different appointment method for corporate tax and this appointment method could affect private R&D. These distinctions are fundamental bases for the following empirical analysis of the effectiveness of the tax credits because constructing the valid comparison group is essential for the quasi-experimental approach.

The first criterion is the base for the credit. While most states calculate the credit based on qualified R&D expenses with the conformity of the federal law, some do this based on employment. The employment-based R&D tax credits can be understood as an effort to increase employees within a state as a primary objective. According to data primarily collected from previous research and then updated and double-checked from the database of the Tax Research Network, 35 states have utilized tax credits based on R&D expenses, while 3 states (i.e. Mississippi, Oklahoma, and Vermont) have offered employment-based tax credits as shown in Table 7. In addition, the remaining 13 states do not offer any type of state R&D tax credit.

Figure 1 represents the geographical distribution of state R&D tax credits by the base for the credit.

Table 7. The different utilizations of state R&D tax credits by the base of the credit

The base of state R&D tax credits	States
R&D Expense	Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Georgia, Hawaii, Idaho, Illinois, Indiana, Iowa, Kansas, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Missouri, Montana, New Jersey, New Mexico, North Carolina, North Dakota, Ohio, Oregon, Pennsylvania, Rhode Island, South Carolina, Texas, Utah, Washington, West Virginia, Wisconsin
Employment	Mississippi, Oklahoma, Vermont
None	Alabama, Alaska, District of Columbia, Florida, Kentucky, Nebraska, Nevada, New Hampshire, New York, South Dakota, Tennessee, Virginia, Wyoming

Source: Sigalla and Viard (1999), State Science and Technology Institute (1997), Rashkin (2003) and updated and double-checked from the database of CCH Tax Research Network

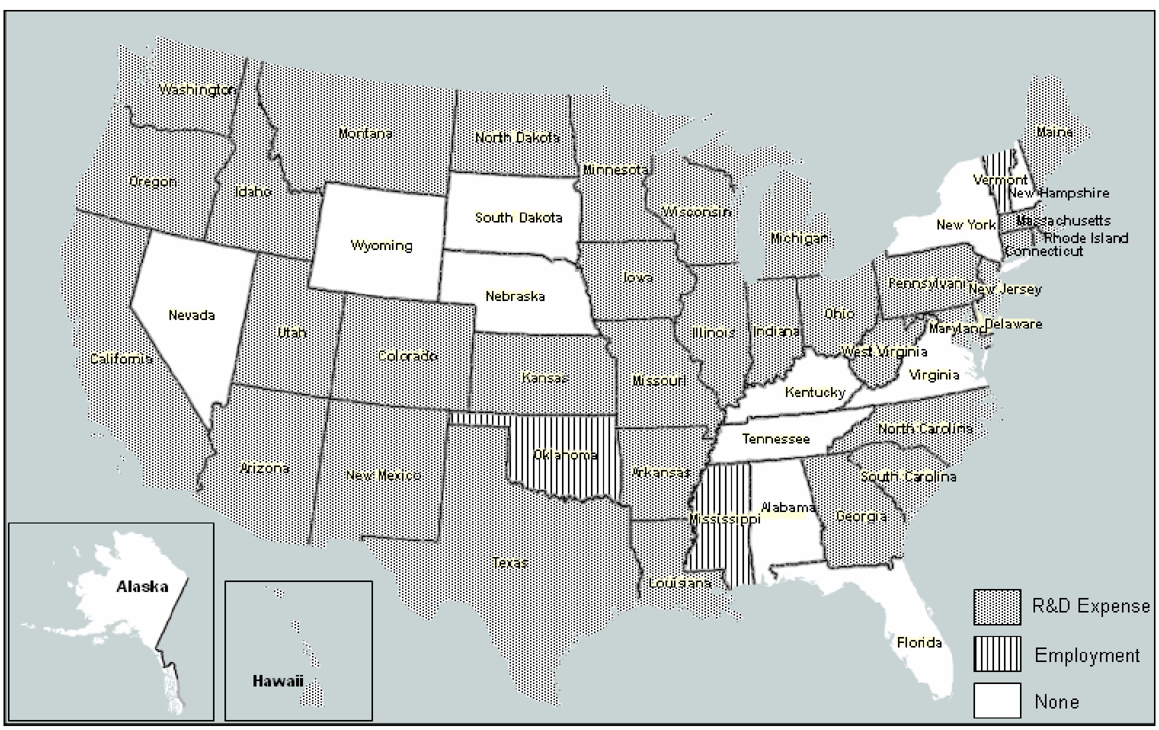


Figure 1. Geographical distribution of state R&D tax credits depending on the credit base

The second criterion is the operation year. The credits were enacted in quite different time periods across states and indeed some are already expired while others are not. In detail, among thirty-eight states having state R&D tax credits, twelve states introduced them during the 1980s, and seventeen states made them in the 1990s, while nine states enacted them during the 2000s. Table 8 shows how state R&D tax credits can be grouped by their enactment years with their bases of the credits and indicates specific operation years in detail. Some state tax credit programs are made as temporary, which means the expired year is specified, while other state programs are made as permanent. Although some state R&D tax credits were already expired, the R&D credit is getting more popular across states. This is one of most fundamental pieces for evaluation, because it determines the time variable that indicates before and after a program. Figure 2 represents the geographical distribution of state R&D tax credits by the operation year.

Table 8. The different utilizations of state R&D tax credits by the enactment year and the base of the credit

The base of the credit / The enactment year	R&D Expense	Employment
1980s	Minnesota (1982-), Iowa (1985-), Wisconsin (1986-), California (1987-), Kansas (1987-2000), North Dakota (1988-), Colorado (1989-), Oregon (1989-2011)	Mississippi (1989-)
1990s	Illinois (1990-2003), Indiana (1990-2013), Massachusetts (1991-), Arizona (1994-2003), Connecticut (1994-), New Jersey (1994-), Rhode Island (1994-), Missouri (1995-2004), Washington (1995-2004), Maine (1996-), North Carolina (1996-2006), Arkansas (1997-), Pennsylvania (1997-2006), Georgia (1998-2004), Montana (1998-2010), Delaware (1999-2006), Utah (1999-2010)	Vermont (1992-1996) Oklahoma (1993-2003)
2000s	Hawaii (2000-2010), Maryland (2000-2006), New Mexico (2000-), Texas (2000-2009), Idaho (2001-2005), Ohio (2001-), South Carolina (2001-), Michigan (2002-), Louisiana (2003-2006), West Virginia (2003-)	

Source: Sigalla and Viard (1999), State Science and Technology Institute (1997), Rashkin (2003) and updated and double-checked from the database of CCH Tax Research Network

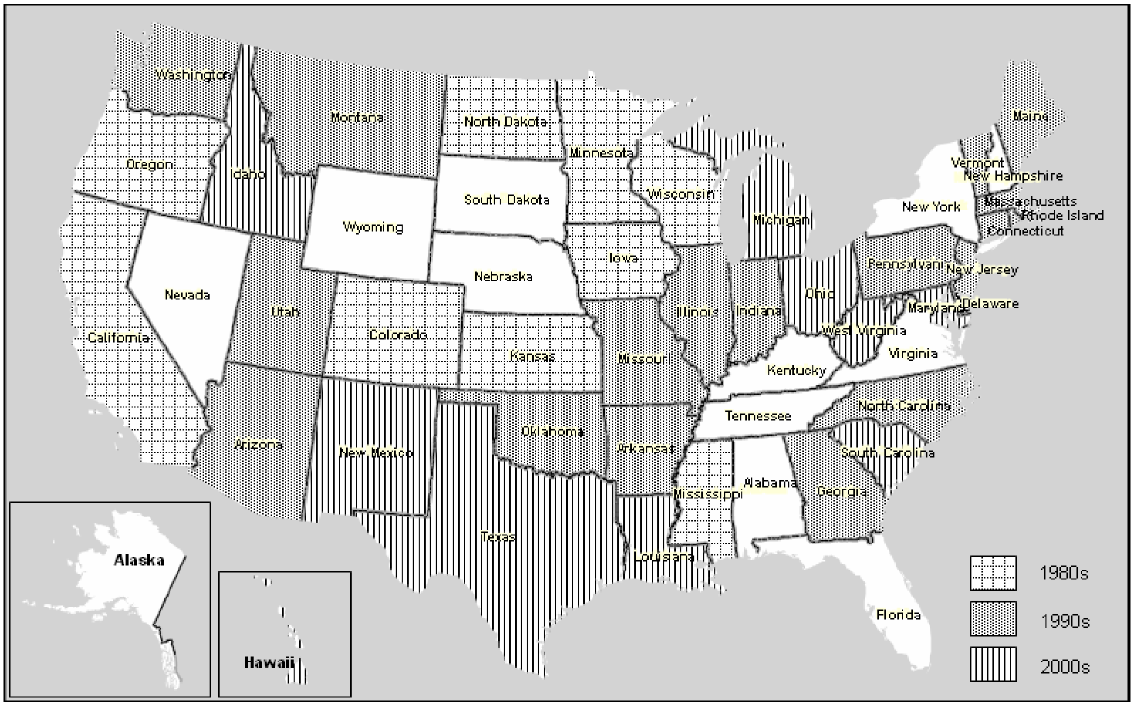


Figure 2. Geographical distribution of state R&D tax credits based on operation year

Regarding to variations of the operation year of the credits, relationships with the level of R&D performance by state provide an interesting distributional characteristic. The geographical distribution of R&D activity within US has been quite skewed, in the sense that six states have performed approximately 50% of whole R&D activity in US and ten states have done nearly 70% of this activity during the 1990s (see for example, NSF, Division of Science Resource Studies, 1995, 1998, 2001, and 2002 and NSF, National Science Board, 2000). However, the adoption of the R&D tax credit law is not consistent with this distribution as represented in Table 9. Among top ten states performing industry R&D activity in the 1990s, only 2 states (i.e. California and Illinois) have offered R&D tax credit since the 1980s, 4 states (i.e. Massachusetts, New Jersey, Pennsylvania, and Washington) enacted R&D tax credit during the 1990s, and other 3 states (i.e. Ohio, Michigan, and Texas) have only short history. In addition, all of the R&D tax credits of these nine states are R&D expense-based. The state of

New York still doesn't offer the state R&D tax credit. Some states with the low level of R&D performance enacted the R&D tax credits much earlier. This fact provides the basis of the better evaluation by enabling to construct the valid control group which means the comparable set of states having tax credit and states not having.

Table 9. The division of state R&D tax credits among top ten states performing private R&D

The base of tax credits	Enactment Year	Top ten states performing private R&D based on the amount of R&D spending in 1990s
R&D-expense	1980s	California (1987-), Illinois (1990-2003)
	1990s	Massachusetts (1991-), New Jersey (1994-), Pennsylvania (1997-2006), Washington (1995-2004)
	2000s	Ohio (2001-), Michigan (2002-), Texas (2000-2009)
None	N/A	New York

Source: Sigalla and Viard (1999), State Science and Technology Institute (1997), Rashkin (2003) and updated and double-checked from the database of CCH Tax Research Network

The third one is the application range of the tax credit. Most state tax credits can be claimed for the every eligible R&D expenditure, while three states allow the credits for only R&D expenditures of specific industries or those within specific areas. These are Arkansas, Colorado, Michigan, and Washington as shown in Table 10. For example, Michigan, as one of top ten states performing private R&D, has offered the credit to only the pharmaceutical industry and Colorado has provided the credit only within enterprise zones. Based on the limited effect of R&D tax credits from their limited application, these three states are categorized by the control group instead of the treatment group in this study. Indeed, while most state laws are incremental tax credits, which means only amount over a certain base level is able to be deducted from corporate income taxes, some states have utilized tax credits as a non-incremental way.[52]

[52] For comparing these two approaches, refer to Hall and Wosinska (1999).

Table 10. The characteristics of targeted/limited state R&D tax credits

States	Operation year	Features
Arkansas	1997-	Limited by industry: the biotechnology industry
Colorado	1989-	Limited by area: the enterprise zone
Michigan	2002-	Limited by industry: the pharmaceutical industry
Washington	1995-2004	Limited by industry: advanced computing, advanced materials, biotechnology, electronic device technology, and environmental technology

Source: Sigalla and Viard (1999), State Science and Technology Institute (1997), Rashkin (2003) and updated and double-checked from the database of CCH Tax Research Network

The fourth one is the state's appointment method for calculating the amount of income for the base of tax. This criterion is not exactly relevant to the R&D tax credit, rather relevant to general state tax structure, but, it is important for calculating the tax amount related to the R&D activity as the following way. Some states follow the federal law, which uses three-factor formulas, evenly distributing corporate income as gross receipts, payroll and property. However, some states utilize only one-factor formulas by using gross receipts, or double-weighting method based on gross receipts. Table 11 summarizes the each state's tax structures with its state R&D tax credit utilization.

Out of thirty-eight states having the state R&D tax credits, eight states use "the equally distributed three-factor formulas" including Delaware, Kansas and Oklahoma, twenty-eight states use "the double- weighting method based on sales" or "the single-factor formulas based on sales", including California, Illinois, New Jersey, and Texas. The remaining two states, namely Montana and Washington, have unique state tax structures. Montana does not impose a general sales and use tax, and therefore there is no possiobe tax benefit to the R&D activity. Washington does not impose a corporate income tax, but instead impose the Business and Occupation (B&O) tax based on gross receipts. Because the B&O tax is entirely based on gross receipts, it is expected that to generate possible tax benefit related to the R&D activity as like single-factor formula or double-weighted formula based on sales.

Table 11. The differences of state tax structures with different utilizations of state R&D tax credits

State tax structure	The base of tax credit	States
Three-factor formula (same as the federal law)	R&D Expense	Delaware (1999-2006), Kansas (1987-2000), North Dakota (1988-), Rhode Island (1994-), Utah (1999-2010)
	Employment	Mississippi (1989-), Oklahoma (1993-2003), Vermont (1992-1996)
	None	Alabama, Alaska, District of Columbia
Single-factor formula or double-weighted formula based on sales (possible tax benefit related to the R&D activity)	R&D Expense	Arkansas (1997-), Arizona (1994-2003), California (1987-), Colorado (1989-), Connecticut (1994-), Georgia (1998-2004), Hawaii (2000-2010), Idaho (2001-2005), Illinois (1990-2003), Indiana (1990-2013), Iowa (1985-), Louisiana (2003-2006), Maine (1996-), Maryland (2000-2006), Massachusetts (1991-), Michigan (2002-), Minnesota (1981-), Missouri (1995-2004), New Jersey (1994-), New Mexico (2000-), North Carolina (1996-2006), Ohio (2001-), Oregon (1989-2011), Pennsylvania (1997-2006), South Carolina (2001-), Texas (2000-2009), West Virginia (2003-), Wisconsin (1986-)
	None	Florida, Kentucky, Nebraska, New Hampshire, New York, Tennessee, Virginia
N/A	R&D Expense	Montana (1998-2010)* Washington (1995-2004)**
None	none	Nevada, South Dakota, Wyoming

Notes
· * Montana does not impose a general sales and use tax, and therefore there is no possible tax benefit to the R&D activity.
· ** Washington does not impose a corporate income tax, but instead impose the Business and Occupation (B&O) tax based on gross receipts. Because the B&O tax is entirely based on gross receipts, it is expected that to generate possible tax benefit related to the R&D activity as like single-factor formula or double-weighted formula based on sales.
Source: Sigalla and Viard (1999), State Science and Technology Institute (1997), Rashkin (2003) and updated and double-checked from the database of CCH Tax Research Network

Among thirteen states not having the state R&D tax credits, three states, including Alabama and Alaska, use “the equally distributed three-factor formulas” and seven states, including Florida and New York, use “the double- weighting method based on sales” or “the single-factor formulas based on sales” (Rashkin, 2003). The remaining three states, including Nevada, South Dakota, and Wyoming, do not have state corporate taxes.

As Rashkin (2003) pointed out, the initial R&D activity might be closely related to building new facility or hiring new researchers rather than increasing sales immediately, and therefore, there is a possible benefit from the state's one-factor formulas or double weighting method based on gross receipts. Therefore, the latter cases of each category are expected to create the tax benefit from the R&D activity by under-appointing property and payroll factor to the amount of income.

Based on the above criteria which indicate the different utilizations of the state R&D tax credits, this study constructs samples and divides them into two groups (i.e. the experimental group and the control group) for empirical analysis of measuring the effectiveness of state R&D tax credits. Only states confirming next three conditions will be classified into the treatment group; (1) states having the R&D expense-based credits and allowing credits to every qualified R&D spending, (2) states enacting credit laws al least until 1999, and (3) states having the state tax system as "single-factor formula or double-weighted formula based on sales". Then, only states confirming next three conditions will be categorized as the control group; (1) states not having tax credits *or* states having the R&D expense-based credits but allowing credits to only limited R&D spending based on the industry and area, (2) states enacting credit laws after 1999, and (3) states having the same state tax system as above. Other states not being satisfying the above criteria are excluded in this study.

Table 12 reorganizes state R&D tax credit programs based on the above selection rules for defining the treatment group and the control group in this study. First of all, all the states including the analysis have a state tax system as "single-factor formula or double-weighted formula based on sales". Among them, the states are sub-grouped by the detailed selection rules mentioned above. Accordingly, the experimental group consists of sixteen states and the control group consists of nineteen states, and their enactment years are quite varied. Again, among the top ten states performing R&D activity in US, six of them are constituted in the experimental group and three of them are consisted of the control group,[53] which provides one of the fundamental bases for an appropriate comparison.

Table 12. The selected states for the analysis of the effectiveness of state R&D tax credits based on different state tax structures and different utilizations of state R&D tax credits

Overall State tax structure and R&D tax credit	Detailed State R&D tax credit	States	Application of this study
States (1) having R&D expense-based tax credits and (2) having state tax systems as single-factor formula or double-weighted formula based on sales	Every qualified expenditure before 2000	Arizona (1994-2003), California (1987-), Connecticut (1994-), Georgia (1998-2004), Illinois (1990-2003), Indiana (1990-2013), Iowa (1985-), Maine (1996-), Massachusetts (1991-), Minnesota (1981-), Missouri (1995-2004), New Jersey (1994-), North Carolina (1996-2006), Oregon (1989-2011), Pennsylvania (1997-2006), Wisconsin (1986-)	Included Treatment group
	Every qualified expenditure but, after 2000	Hawaii (2000-2010), Idaho (2001-2005), Louisiana (2003-2006), Maryland (2000-2006), New Mexico (2000-), Ohio (2001-), South Carolina (2001-), Texas (2000-2009), West Virginia (2003-)	Included Control group
	Limited expenditure before 2000	Arkansas (1997-), Colorado (1989-)	
	Limited expenditure after 2000	Michigan (2002-)	
States (1) not having R&D tax credits and (2) having same state tax systems as the above	N/A	Florida, Kentucky, Nebraska, New Hampshire, New York, Tennessee, Virginia	

In this section, I analyzed the different utilizations of state R&D tax credit programs by using four criteria, which are (1) the base for the credit, (2) the operation year, (3) the application range, and (4) the state appointment method for corporate tax. Then, thirty-five states were selected for analyzing the effectiveness of state R&D tax credits based on the comparability of their state R&D tax credit programs.

[53] Washington state is excluded from the analysis because of its unique state tax structure and its limited utilization of state R&D tax credit program.

In the next section, I examine state R&D tax credit programs of selected five states in detail for understanding this program comprehensively and specifically. The five states are chosen among top ten states performing private R&D based on their R&D spending in the 1990s and their R&D tax credit programs are similar in some aspects but distinguished in some aspects.

4.3.2 The detailed features for the selected states policies

As mentioned above, the state R&D tax credit programs are quite different across states. Based on the level of R&D spending, which tends to be geographically concentrated and skewed in the US, five states' tax credit laws (i.e. California, Texas, Massachusetts, Michigan and Washington) are investigated in detail. These states are selected because they have been ranked within top ten states of the industry R&D expenditures and also characterized by having similar R&D tax credit policies, which are the R&D expense-based credits as the modifications of the federal system, but their utilizations of tax credits are different somewhat as shown in Table 13. I examine different backgrounds and different specifications of tax credit program of each selected state in detail.

Table 13. The selected five states' R&D tax credit utilizations among top ten states of the industry R&D expenditures in the 1990s

The applicable R&D expenditure for tax credits	Enactment year	Top ten states performing private R&D based on the amount of R&D spending in 1990s	
		selected	unselected
Every qualified R&D expenditures	1980s	California (1987-)	Illinois (1990-2003)
	1990s	Massachusetts (1991-) Texas (2000-2009)	New Jersey (1994-) Pennsylvania (1997-2006)
	2000s		Ohio (2001-)
Limited R&D expenditures	1990s	Washington (1995-2004)	
	2000s	Michigan (2002-)	
None	N/A		New York

4.3.2.1 California

California enacted its state R&D tax credit in 1987. The US introduced the R&D tax credit at the federal level in 1981 and only some states had these tax credits since the 1980s. So, California is one of the pioneer states adopting state R&D tax credit programs. We can understand this earlier enactment as the California's strong necessity from its biggest R&D performance among states especially in private R&D activity. California also has utilized a variety of state tax credits, compared with other states. For example, California has developed manufacturer investment credit as well. The various tax incentives and industrial policy measures are aimed to attract private firms within California, which has a relatively high tax environment (Hall and Wosinska, 1999).

The definitions of qualified research expense and base year for calculating credit as an incremental way are tied to the federal law. The credit rate started at 12% for qualified research expenses and 8% for basic research payments, and then credit rates had increased several times. In the year of 2000, the credit rate reached 24% for qualified research expenses and 12% for basic research payments, in conformance with the federal law. With increases in credit rates, California improved this policy in many ways, such as the consideration of small/start-up companies by using the three-tiered approach, in other words, the Alternative Incremental Credit (AIC) since 1997, which also conforms to the federal law. California also made this law permanent based on the theoretical recommendation, which tells us the only permanent tax credit affects private R&D investments because individual firms make a decision for R&D investments as a long term. Compared with the federal R&D tax credit law which has been a temporary law, California's law was made permanent in 1992. Indeed, California is one of the states which allow a private income tax credit as well as a corporate income credit.

Based on the above analysis, California is one of few states having a well-developed R&D tax credit policy with the largest private R&D performance among US states. Regarding to the federal government supported R&D, direct federal R&D funding has declined in terms of the distributional share among states, while the FFRDCs have been widely and persistently operated in California. In 1985 federal funding consisted of 66% of R&D performed in California with 46% of R&D performed in US but in 1997 these portions decreased by 30% and 28% respectively (California Council on Science and Technology, 1999). Importantly California had ranked one of top states receiving federal funding during 1980s and 1990s (Massachusetts Technology Collaborative, 2004). Although private R&D investment has taken the larger portion of total R&D investment and government R&D funding has relatively decreased in the terms of its share of the total R&D spending growth at the whole national level, federal level R&D funding still plays an important role in the R&D activities for the future economic growth. Therefore, attracting federal level R&D funding is another important policy issue for the state of California.

4.3.2.2 Massachusetts

As one of the R&D-intensive states, Massachusetts enacted its R&D tax credit relatively earlier than other states in 1991. The Massachusetts R&D tax credit mostly conforms to the federal law and only difference is the credit rate (i.e. 10% of qualified research expense and 15% of basic research payment, instead of 24% and 12% respectively). The state of Massachusetts applies a separate rule to start-up companies with the conformity of the federal law and defense-related R&D activities. The particular consideration of defense-related R&D activities can be understood in the context of the higher portion of defense-related R&D activities within this state.

In sum, Massachusetts has developed the state R&D tax credit for stimulating private R&D activities within its territory.

Along with its strong industry R&D activities, Massachusetts also has ranked high in federal R&D funding. Even though the portion of federal funding has been getting low with the decline of defense-related R&D activities, the state of Massachusetts has been an important recipient of federal R&D funding related to the extensive university R&D funding. Accordingly, Massachusetts has also focused on establishing relationships between universities and industries for enriching R&D activities and their commercialization. This balanced development of supporting private R&D activities indirectly through tax credits and directly through attracting federal level R&D funding and reinforcing its utilization is one of the strengths of R&D policies of Massachusetts. However, securing federal R&D funding has been more competitive and it tends to disperse across states rather than concentrated on a few states. Massachusetts Technology Collaborative (2002) figured out that there was the overall loss of market share of the federal R&D funding in the state based on the above facts and there were two major transitions of its allocation, namely, the continued increase in health-related industries and the substantial decrease in defense-related industries. Then, it (2004) also figured out that Massachusetts repositioned as a major beneficiary of federal funding for not only health-related industries but also defense-related industries. Indeed, the more states has utilized and developed the state R&D tax credits for attracting private firms. To cope with the severe regional competition, the more intensive strategies are required for every state including Massachusetts.

4.3.2.3 Michigan

Even though Michigan has been historically the second highest private R&D performance state based on its private R&D spending,[54] its R&D tax credit program is in its initial stage. Namely, Michigan's state R&D tax credit just began in 2002 and it is applied only to pharmaceutical companies. The eligible expenditure is qualified research expense defined as the federal law and the credit rate is 15% that is lower than the federal one. Indeed, there is no credit for basic research payment. The Michigan R&D tax credit has been operated as temporary and will be expired in 2009.

One possible explanation for this limited R&D tax incentive program is that Michigan has established a business-friendly environment with a generous tax policy, which means its corporate tax rate is as low as 1.9%, compared with California (8.84%) and Massachusetts (9.5%). In other aspects, Michigan's R&D policy has focused more on direct government funding associated with universities. As Michigan has been characterized by high R&D performance of universities with industries, the state policy has focused on encouraging commercialization of university R&D, supporting spin-off or start-up companies and capitalizing venture capital rather than using indirect incentives such as a tax credit. In this context, Michigan Economic Development Corporation (2002) suggested the R&D tax credit is linked with start-up companies, spin-off companies and universities.

Another possible explanation for the Michigan's limited use of the R&D tax credit, despite its second highest ranking for industrial R&D performance, is the skewed dependency on the R&D activities of the large automobile and other related manufacturers. These highly

[54] Michigan has mostly ranked as the second state performing industry R&D over time (see, for example, NSF, Division of Science Resource Studies, 1995, 1998, 2001, and 2002 and NSF, National Science Board, 2000).

agglomerated and concentrated industries are assumed to undertake large R&D activities in the absence of the tax credit (Secretaries of Technology and Commerce and Trade of Virginia State, 2000).

4.3.2.4 Texas

Texas enacted the R&D tax credit in 2000. It had debates about the efficiency of the state R&D tax credit,[55] and about the legitimacy of the state R&D tax credit. Some insisted on the possibility of its low effect because Texas already had built a good business environment including a low corporate income tax, which is 4.5%. Indeed, some recommended non-incremental credit rather than incremental credit for avoiding an inefficient link between past and current R&D investments. There was other suggestion to use a sales tax exemption rather than tax credit due to Texas' relatively high sales tax. There was also a question about an uncertain geographical boundary of social returns from increasing R&D, in the sense that spillovers could occur across states. In spite of the above notions, the efforts for supporting the state R&D tax credit, which is mostly made by "the Technology Business Council," finally made it possible to pass this law in 2000 along with other economic development credits.

Basically the eligible R&D expenditures are qualified research expenses and basic research expenses that conform to the federal law, and the credit rate is applied as 5% for both expenses, which is lower than the federal law as 20%. Texas' R&D tax credit is also utilized as an incremental way, which means only exceeding amount to a base is eligible for the credit, as same as the federal law.

[55] For the details, refer to Sigalla and Viard (1999).

For better implementation, Texas utilizes a bonus system linked with area development, which is called as the "strategic investment area." In addition, it applies the Alternative Incremental Credit (AIC), as an alternative method of calculating the base amount, for small new business companies, because these companies are mostly featured by high R&D spending and rapid sales growths and therefore cannot receive the credit under the general rule. These two features can be understood as Texas' efforts for increasing effect of R&D tax credit through the better utilizations.

4.3.2.5 Washington

In 1994, Washington adopted the state R&D credit against the Business and Occupation Tax, which is a gross receipts tax as an alternative tax to a corporate income tax. The credit rate varies depending on the type of business (i.e. nonprofit and proprietary), unlike other states that apply different credit rates for basic research payments and qualified research expenses. The credit rate for nonprofit corporations and associations is 0.484% and that for proprietary businesses is 1.5% in 2004. Those were previously 0.515% and 2.5%, respectively. It means that the rates of the Washington R&D tax credit were going down although other states tended to increase credit rates. As a different form for encouraging R&D, Washington temporarily applied the credit of $1,000 per new employee in manufacturing and R&D companies in distressed areas from 1986 to 1988.

In addition, Washington is one of the states that use a targeted R&D tax credit, which means only qualified industries can apply the credit. Qualified industries are advanced computing, advanced materials, biotechnology, electronic device technology, and environmental technology, based on the different potential effect of each industry.[56] Another difference is the non-incremental approach, which means that every expenditure related to research, not only the increased portion over the base-year spending, can be claimed. In this sense, the state of Washington has utilized the R&D tax credit in interesting ways along with its distinguished state tax system.

4.3.2.6 Overall comparisons

Table 14 summarizes different features of selected states' R&D tax credits. First, the enactment years are quite varied even though the selected states are R&D intensive states in US as mentioned above. Whether the credit is permanent or temporary is another important issue. Some argued it should be utilized permanently because R&D investment tends to be determined and performed in the long-term basis and therefore, the temporarily utilized and short-term based R&D credits could not properly affect to increase R&D spending. However, some focused on the unclear effect of tax credits and therefore, insisted on the temporary operation of this incentive. Indeed, whether it is incremental or non-incremental indicates the difference of applicable R&D expenses. Following the federal law, most states' programs are utilized in the incremental way, which means only the expenses over base amount can be eligible for claiming credit. However, some have argued that the incremental credit is unreasonable because it makes an inadequate linkage between previous R&D investments and current ones. The application ranges and credit rate are also different based on their own specific needs. The state of Washington has utilized the most unique R&D tax credit program. The adoption of the

[56] The source of industry selection is "Incentives for High Technology" by the Research and Legislation and Policy Divisions of the Department.

Alternative Incremental Credit (AIC) is also important in the sense that the AIC is specially targeted and supports to small start-up companies that experienced rapid sales growths with high R&D spending.

In this section, I examined the selected state R&D tax credit programs in detail. Each state R&D tax credit program has been implemented in a unique way rather than a uniform way, based on each state's different R&D environment, business environment, and tax environment.

In the next section, I discuss the previous evaluations of state R&D tax credit programs. Then, I also discuss the previous evaluations of the relevant R&D policies, which are the federal R&D tax credit and direct funding.

Table 14. Detailed comparisons of selected state R&D tax credit programs

	California	Texas	Washington	Michigan	Massachusetts
Enactment year	1987 (claimed from 1988)	2000	1995	2002	1991
Modification	1993, 1996, 1998, 2000	2002	1998		1992, 1998
Permanent *vs* temporary	Permanent (from 1993)	Temporary (expires in 2009)		Temporary (expires in 2009)	Permanent (from 1992)
Incremental *vs* non-incremental	Incremental	Incremental	Non-incremental	Incremental	Incremental
Targeted *vs* general	General	General	Targeted	Targeted	General
Corporate tax rate	8.84%	4.5%	n/a	1.9%	9.5%
Credit rate	15% qualified, 24% basic research	5% qualified, basic research	0.484% nonprofit, 1.5% proprietary	6.5% qualified research	10% qualified, 15% basic research
Alternative Incremental Credit (AIC)	Yes	Yes	No	No	No

Notes:
· California also allows the R&D tax credit to individual income tax.
· In Massachusetts, defense related activities are treated separately.
· In Texas, the bonus credit is allowed for the R&D activity performed within “strategic investment area”.

4.4 PREVIOUS EVALUATIONS OF STATE R&D TAX CREDIT PROGRAMS AND OTHER RELEVANT R&D POLICIES

In the first part of this section, I review previous evaluations of state R&D tax credit programs. The available studies are undertaken for individual state R&D tax credit program separately not for a number of state programs together, and are also quite limited. One of them is a study on California R&D tax credit by Hall and Wosinska (1999). Second one is a study on Virginia R&D tax credit, conducted by Secretaries of Technology and Commerce and Trade of Virginia State (2000). Third ones are three consecutive studies on Washington R&D tax credit by the Department of Revenue of Washington State in 1997, 2000, and 2003. These studies pointed out some significant issues for the program evaluations despite not providing any systematic statistical analysis.

Unlike the limited number of studies evaluating state R&D tax credits, there are numerous empirical studies which perform the evaluations of other government R&D policies. In the second part of this section, I examine previous evaluations of the federal R&D tax credit, as the most relevant R&D policy in the ground that state R&D tax credits are mostly modeled on the federal R&D tax credit. Compared to the state-level analysis, the federal-level analysis is well-developed in terms of the methodology and the frequency. The reliable empirical results are provided from various researchers and based on them, there is a general consensus of the positive effect of this program even though the results are somewhat different in terms of the degree of sensitivity (See, for example, Cordes, 1989; OTA, 1995; Hall and Van Reenen, 2000).

In the third part of this section, I examine the evaluations of direct funding as a complementary R&D policy with indirect R&D policies including tax credits. As like the evaluations of the federal R&D tax credit, the evaluations of these R&D policies are widely available and conducted by using various approaches and datasets. These empirical tests of relationship between public R&D spending and private R&D spending for finding the effect of direct interventions could provide another important insight for evaluating state R&D tax credits as well. The possible effects of direct funding are a stimulating effect as a positive effect or a crowding out effect as a negative effect, and the findings are some mixed among various empirical analyses.

These previous studies provide the possibility of statistical evaluation approaches which can be applicable to empirical specification for evaluating state R&D tax credit programs and the extensive sets of significant factors which affect to private R&D.

4.4.1 Previous evaluations of state R&D tax credits

First, I assess the previous evaluations of state R&D tax credits. Although many states have eagerly adopted the R&D tax credits for reinforcing economic development based on the theoretical justification of the essential role of R&D previously mentioned here, few evaluations have been conducted. Broadly, economic development strategies have been rarely evaluated in spite of their broad utilizations with enthusiastic efforts. Among few, the notable examples are Hall and Wosinska's analysis (1999) for California's R&D tax credit and two state government-performed analyses for Virginia and Washington R&D tax credits.

First, Hall and Wosinska (1999) evaluated the R&D tax credit law of California. Even though they did not provide empirical analysis, they suggested two possible theoretical approaches for evaluating tax credit in general. The first possible approach is testing whether a tax policy encourages a social return to be equal to a social cost. Hall and Wosinska (1999, p.16) suggested "comparing the marginal return to industrial R&D dollars at the societal level to the opportunity cost of the extra tax dollars in another way." However, this is too theoretical approach and there is no possible way to directly measure a social return and a social cost, which makes this kind of evaluation hard in practice.

The second possible approach is a cost-benefit analysis. Here, the cost can be measured by the loss of tax revenue and benefit is indicated by the increase of R&D spending. The result can be interpreted as tax credit is better than direct funding or vice versa. In other words, if the increased R&D spending is more than the subsidy, the tax credit is the more efficient way to encourage private R&D than direct funding. From the cost-benefit approach, it is hard to evaluate whether the increase of total R&D spending (private returns) exceeds to the social returns or not because it is not mostly unknown what the optimal R&D level in a society is. If private returns are higher than social returns, it means overinvestment. Fortunately, as previous works revealed the gap of private and social returns is substantially high, the overinvestment is rare in real world and therefore cost-benefit analysis can be meaningful.[57] However, they also mentioned the difficulty of this kind of analysis at the state level due to data availability, but it is possible at the federal-level analysis.

Alternatively, Hall and Wosinska (1999) suggested a comparative analysis of state tax programs. However, this kind of comparative study requires time-consuming works for collecting the historical data of individual state tax legislation. They showed us the complexity of the tax environment related to the state R&D tax credit by the interaction with the federal level R&D tax credit and various forms of credits across states. They assumed the positive effect of R&D tax credits based on the theoretical justification from positive externalities of innovative

[57] The previous works include Griliches (1992), Mansfield (1965) and Bernstein and Nadiri (1988).

activities and localized R&D spillovers. They also indicated a possibility of a zero-sum game if every state provides a R&D tax credit with a perspective of the location and allocation of R&D-intensive firms under inter-state competition. Namely, it is expected that the effects on firm location within a state are getting smaller with popular utilizations of state R&D tax credit programs. In conclusion, they mentioned the lack of studies and evaluations of the state level R&D tax credits and subsequently the prompt necessity of them.

Second, Virginia also made an effort to evaluate the state R&D tax credit (Secretaries of Technology and Commerce and Trade of Virginia State, 2000). In this study, it is assumed that the state R&D tax credits can be beneficial by generating spillover effects within the state and attracting firms to the state. For avoiding a zero-sum game, the spillover effect is crucial, but it is still not verified. It is also pointed out the lack of study of the state R&D tax credit programs with mentioning that the evidences from the federal R&D tax credit are not enough to justify the effects of state-level R&D tax credit programs. As the other aspect, it is questioned that the small amount of tax credit (approximately 1% of total R&D spending) could encourage private firms to perform the large scale of research. It is also mentioned that, compared to the job creation tax credit, the R&D tax credit seems less effective for job creation. From interviews with policy makers and practitioners, it is revealed that the state R&D tax credit is considered as a fundamental and indispensable component of the overall high-technology-reinforcing policy. It is also recommended that the R&D tax credit should be specially targeted to start-up companies and be utilized in order to encourage connections between industries and universities.

Third, the Department of Revenue of Washington State performed three consecutive analyses on the effect of Washington state R&D tax credit in 1997, 2000 and 2003.[58] It

[58] See http://dor.wa.gov/content/Statistical_Reports/stats_HighTechRandD_report.asp for the full report.

measured the economic impact broadly as job creation, R&D investment, the firm's location and state economy and so on. In 1997, based on the survey to the participant companies, it concluded the positive effect of R&D credit on job creation based on the increase in the number of employment and wage. It also found the growth of company income and R&D spending at the industry level. In 2000, it found the growth of share of the selected industry in US and increasing R&D activities in terms of the number of patents and new products. In 2003, by using the same method, it found that the increase in R&D spending and employment yet no clear evidence of job creation and the diversity of economy. In summary, three consecutive analyses told us the positive effect of the program overall.

However, these analyses are based on the simple comparison of economic indicators before and after program without any other controls. So, the above results can hardly capture the pure effect of tax credit because there is no consideration the possible change from other relevant variables without the credits. In other words, the firms could increase employment or income without the tax credit and the above analyses do not take into account this important counterfactual. The simple fact of increasing number of employment or income after adopting the program doesn't guarantee the success of the program.

In sum, the above previous evaluations do not provide the statistically tested empirical evidences of the program effect rather, these analyses relied on descriptive and theoretical recommendations, interviews and survey approaches, or less reliable simple comparison of outcome changes. Therefore it implies that better evaluations are promptly required for this specific regional economic development policy field.

In the next section, I examine the relevant literatures evaluating federal R&D tax credit and these literatures provide possible statistical evaluation approaches which can be referred to for building an empirical specification for evaluating state R&D tax credit programs.

4.4.2 Previous evaluations of the federal R&D tax credit

In this section, I assess the evaluations of the federal R&D tax credit. The federal R&D tax credit program is quite relevant to state R&D tax credit programs as a prototype of most state programs and therefore, the analyses of the effectiveness of the federal R&D tax credit are also quite relevant to the analyses for state programs. For testing the effectiveness of the federal R&D tax credit, the researchers should answer the question: "how much more R&D did firms do given the existence of a tax credit than they would have done if there had been no credit" (Hall and Van Reenen, 2000, p. 458). For doing this, there are two possible approaches. One approach is the user cost of the R&D approach, which is estimating the elasticity of private R&D to a price variable, in other words, the user cost of the R&D.[59] If the user cost of R&D includes a measure of subsidy from tax credit, this approach can directly measure the elasticity of private R&D to tax credits. If it doesn't include a measure of subsidy, the elasticity can be measured by multiplying the measured elasticity by the marginal R&D tax credit rate.[60] This approach is the most popular method. The other approach is the forecasting approach, based on the assumption of that the current R&D spending is a function of the lagged R&D spending, the lagged output,

[59] For the method of measuring the use cost of R&D, refer to Hall and Van Reenen (2000, p. 467-468).

[60] For the detailed explanation and possible variations of the user cost approach, refer to Hall and Van Reenen (2000) and Bloom, Griffith and Van Reenen (2002). Hall and Van Reenen (2000, p. 459) called this approach as "quasi-experiments as price elasticity estimation" as well.

or the expected demand with the dummy variable indicating tax credit, as the alternative one.[61] Within these approaches the selected covariates for controlling relevant effects are usually the lagged sales, the lagged R&D, the current and lagged outputs, and the cash flow. Then, the social benefit-cost ratio, which is the additional private R&D expenditure due to the tax credit relative to foregone tax revenues, can be calculated by comparing the above estimates and tax revenue loss.

Regarding the findings, the earlier studies showed negative or zero effects, however, the recent studies mostly found positive effects, namely the private firms' positive increases in their R&D spending from given tax credit amounts (see, for example, Cordes, 1989; OTA, 1995a; Hall and Van Reenen, 2000). Hall and Van Reenen (2000) pointed out the earlier works are mostly based on the internal tax return data, surveys, and interviews, while later works are based on the statistical methods. Indeed, the degree of sensitivity varies across studies. Even though it is difficult to explain these differences in sensitivity, there is a consensus that the positive effect of the federal level tax credit is generally an one dollar increase in private R&D investment to an one dollar lost in tax, which indicates the low response of the private sector (see, for example, OTA, 1995a and Hall and Van Reenen, 2000).

Next, each study evaluating the federal tax credit is investigated in detail, based on four different kinds of dependent variables. First possible dependent variable is the corporate income tax return data, which was mostly used in the earlier works. By using this type of data, Altshuler (1989) found almost zero effect, namely no increase in R&D spending from credits, while Cordes (1989) found positive increases in R&D spending which were limited to the non-

[61] Hall and Van Reenen (2000, p. 458) called this approach as "natural experiments as R&D demand equation with a shift parameter for the credit" as well.

manufacturing sector. Second, from the industry-level R&D spending data, the most estimates showed significant and positive increases in private R&D spending (for example, Baily and Lawrence, 1987 and 1992). Recently, Mamuneas and Nadiri (1996) also found significant effects of the federal R&D tax credit by using the industry-level data and the user cost approach. Third possible dependent variable is the firm-level R&D spending data, which is provided from Compustat and the most broadly used one. Eisner, Steven, and Sullivan (1983) used the yearly change in firm's R&D spending for detecting each firm's eligibility of receiving tax credits because the R&D tax credits are determined based on the difference between the current and previous R&D spending, and they found no clear evidence. However, the later works, such as Berger (1993), Hall (1993), and Hines (1993), which used the user cost approach, found a positive effect of the federal tax credit, with approximately one additional dollar increase in private R&D spending to one dollar lost in tax. Fourth, by using the country-level R&D spending data, Bloom, Griffith and Van Reenen (2002) found the amount of private R&D spending increased with the increase in the amount of tax credit.

There are also studies of the effects of R&D tax credits for other countries (for example, Mansfield and Switzer, 1985; Bloom, Griffith and Van Reenen, 2002; Leyden and Link, 1993) and some firm level surveys (for example, Mansfield, 1986; Eisner, Albert and Sullivan, 1984). Cordes (1989) pointed out that these analyses provide the less clear evidences due to the low level of significance of estimates compared to the above studies. Meanwhile, there is a study examining the effect of tax credit depending on firm size. Koga (2003) found the larger effect of tax incentives on R&D spending for large firms by using data of Japanese firms and the user cost approach. In his study, firm size is defined by total amount of capital. The selected previous studies on evaluating federal R&D tax credit are demonstrated in Table 15.

Table 15. The summary of previous studies on the federal R&D tax credit

Type of observational Unit	Studies	Econometric approaches	Dependent variables	Control variables	Results (Elasticity/ BC ratio)
Corporate income tax return data	Altshuler (1989)	NA	NA	NA	No effect
	Cordes (1989)	NA	NA	NA	Positive
Firm	Eisner, Steven and Sullivan (1983)	Forecasting	R&D	Lagged R&D, current and lagged sales	Insignificant
	Mansfield (1986)	Survey	NA	NA	Positive
	Swenson (1992)	Forecasting	Log R&D	Sales, change in long-term debt, dummy for industry	Positive (NA)
	Berger (1993)	Forecasting	R&D	R&D/sales (firm/ind.), investment/sales(firm/ind.), CF/sales, GNP, Tobin's q	Positive (1.0-1.5/1.74)
	McCutchen (1993)	Forecasting	Research intensity	Past NCFs, diversity, CF/sales, % drug sales	Positive (0.28-1.0/ 0.29-0.35)
	Hall (1993)	R&D User cost	Log R&D	Lagged R&D, current and lagged output	Positive (1.0-1.5/2.00)
	Hines (1993)	R&D User cost	R&D	Sales & Tax price (domestic/foreign), dummy for industry and firm	Positive (1.2-1.6/ 1.3-2.0)
Industry	Baily and Lawrence (1987, 1992)	Forecasting	Log R&D	Lagged R&D, current and lagged output	Positive (0.75/1.30)
	Mamuneas & Nadiri (1996)	R&D User cost	R&D	Output, public R&D, factor prices	Positive (0.95-1/0.95)
	Bloom, Griffith and Van Reenen (2002)	R&D User cost	Log R&D	Output, dummy for time and country	Positive (0.14/NA)

Source: Cordes (1989), OTA (1995a), Hall and Van Reenen (2000), updated by author.

The overall finding, which is approximately one dollar increase of R&D spending with one dollar lost in tax revenue, tells us the low response of the federal R&D tax credit. This result can partially come from the low response especially in the earlier operation years, which is revealed in the earlier works in which the negative or zero effects were found (Hall and Van Reenen, 2000). Accordingly, there is a room for finding the larger positive effect by using only the recent datasets or separating the effects by several time periods. Or, this result can partially come from less efficient characteristics of the federal R&D tax credit such as operating as a temporary law. In the next section, I discuss the evaluations of direct public funding.

4.4.3 Previous evaluations of direct public funding

Next, I assess the evaluations of direct public funding. Direct public funding as one type of direct government intervention is closely related to R&D tax credits as one type of indirect government intervention, in that they have complementary relationships by playing different roles in affecting private R&D.

For testing the effectiveness of direct public R&D funding, empirical analyses mostly applied regressions of private R&D funding on public R&D funding with some control variables. Positive coefficients in public R&D funding (i.e. the increase in private R&D funding with the increase in public R&D funding) indicate stimulating effects, while negative coefficients (i.e. the decrease in private R&D funding with the increase in public R&D funding) indicate crowding out effects. Stimulating effects means public R&D funding could generate the increase of private R&D funding as a complementary relationship and ultimately result in the productivity growth. Crowding out effects means public R&D investment simply substitutes for private R&D investment in the sense that without public R&D investments, private firms might make these investments at their own expense and therefore, with public R&D investments, private firms might decrease their own R&D investments.

Evaluations testing the relationship between private R&D funding and public R&D funding, which are either complements as positive effects or substitutes as negative effects, have been performed by numerous researchers and by using a variety of methods and types of observational units. David, Hall and Toole (2000) recognized these methodological differences by four categories. Types of observational units are categorized by the laboratory, the firm, the industry, and the aggregate form. Econometric methods are divided by the cross-section, the panel, and the macro level, and the instrumental variables. From these different approaches, researchers generally concluded that there is a positive effect of public R&D funding on private R&D, indicating the complementary relationship, even though there are some exceptions which find the crowding out effect, or substitution effects. In these previous empirical analyses, dependent variables are usually the amount of private R&D expenditure or number of employees and independent variables are usually the amount of government R&D expenditure or contracts. These variables are constructed with some different formats, such as the log, the change, or the ratio. The summary of previous studies is described in Table 16.

From the multiple methods and multiple datasets which was explained above, the analyses generally show the positive effect of direct public funding, which is called stimulating effects. It tells us that direct public funding has complementary relationship with private R&D, in other words, promoting private R&D.

Table 16. The summary of previous studies on government direct intervention

Econometric approaches	Studies	Type of observational unit	Dependent variables	Independent variable	Control variables	Results
Cross-section	Scott (1984)	Laboratory	Log (amount of private R&D)	Log (amount of government R&D)	Size, *Dummy* (firm or industry)	Positive (.06-.08)
	Leyden and Link (1991)	Laboratory	amount of private lab budget	amount of government R&D funding to lab	*R&D/Sales*, *Dummy(*basicR, K-sharing, chem/bio)	Positive (.336)
	Link (1982)	Firm	Amount of private R&D/sales	Amount of government R&D/sales	Profit/S, diversification, C4, *Dummy* (governance)	Positive
	Lichtenberg (1984)	Firm	Change in amount of private R&D/sales	Change in amount of government R&D/sales	Size	Negative
	Antonelli (1989)	Firm	Amount of private R&D, Log (private R&D)	Log(Government R&D/total R&D), % Government R&D	Size, profit, share of foreign sales, *Dummy* (diversification)	Positive (.31-.37)
	Busom (2000)	Firm	Amount of private R&D, R&D per employee	*Dummy* (subsidy program participation)	Size, patents, export share, *Dummy* (industry)	Positive (.2)
	Globerman (1973)	Industry	R&D employees / total employees	Amount of government R&D/sales	% of foreign, sales growth, *Dummy* (tech opportunity)	Positive
	Buxton (1975)	Industry	Amount of private R&D/gross output	Amount of government R&D/gross output	C4, diversification, entry barriers	Positive
Panel analysis	Klette and Moen (1998)	Laboratory	Log (amount of private R&D)	Log (amount of government R&D)	Sales, sales sq., cash flow, *Dummy* (time)	Positive (.06)
	Howe & McFetridge (1976)	Firm	Amount of private R&D	Amount of government R&D Grants	Size, profit, depreciation, HHI, *Dummy* (foreign)	Mixed
	Lichtenberg (1987)	Firm	Amount of private R&D	Amount of government R&D	*Dummy* (year), size, sales to government	Insignifi-cant
	Lichtengerg (1984)	Industry	Change in amount of private R&D	Change in amount of government R&D	*Dummy* (industry and year)	Insignifi-cant
	Levin and Reiss (1984)	Industry	Amount of private R&D/production costs	Amount of government R&D/shipments	*Dummy* (technology), basic R&D share, industry age, HHI	Positive

Table 16. The summary of previous studies on government direct intervention (continued)

Econometric approaches	Studies	Type of observational unit	Dependent variables	Independent variable	Control variables	Results
Panel analysis	Levy (1990)	Aggregate macro-level	Amount of private R&D	Amount of government R&D contracts to industry	GDP, private R&D in Europe and Japan, *Dummy* (country)	Positive
	Von Tumzelmann & Martin (1998)	Aggregate macro-level	Change in amount of private R&D	Change in amount of public R&D	Levels of private and public R&D, *Dummy* (country)	Positive
Time-series analysis	Levy and Terleckyj (1983)	Aggregate macro-level	Amount of private R&D	Amount of government R&D contracts to industry	Lagged output & taxes, age, unemployment, R&D stock, government R&D, Reimbursement	Positive
	Lichtenberg (1987)	Aggregate macro-level	Amount of private R&D	Amount of government R&D contracts to industry	Sales, sales to government	Insignifi-cant (.045)
	Diamond (1998)	Aggregate macro-level	Amount of private basic research	Amount of federal basic Research	GDP, time trend	Positive (1.04)
Instrumental variables/first difference	Lichtenberg (1988)	Firm	Amount of private R&D	Amount of government R&D	Sales to government, Size, *Dummy* (year)	Mixed
	Toivanen & Niininen (1998)	Firm	Amount of private R&D	Amount of government R&D	Current and lagged investment, cash flow and interest rate	Negative

Source: David, Hall and Toole (2000)

In this chapter, I examined practical implementations of state R&D tax credits. Many states have implemented these tax credit programs for encouraging private R&D in different ways rather than in uniform ways. I compared state R&D tax credit programs by the base of credit, the enactment year, the application range, and the overall state tax structure. Then, I examined R&D tax credit programs of selected states, which are California, Massachusetts, Michigan, Texas, and Washington, among top ten states performing private R&D, in detail. These states have developed their own state R&D tax credit programs under their unique economic contexts. Then, I discussed previous evaluations of these tax credit programs. These analyses are just a few and rely on descriptive and theoretical recommendations, interviews and survey approaches, or less reliable simple comparison of outcome changes. Compared to few efforts to evaluate state R&D tax credit programs, there are extensive studies on evaluating the federal R&D tax credit and direct funding. These previous studies provide the possibility of statistical evaluation approaches, which can be applicable to empirical specification for evaluating state R&D tax credit programs and the extensive sets of significant factors which affect to private R&D, which are also important for assessing state R&D tax credit programs.

In the next chapter, I discuss the methodology to be used in this study. The most appropriate research design is the quasi-experimental design because outcome differences between states having tax credits and states not having tax credits, which are non-equivalent, will be compared in order to evaluate state R&D tax credit programs as one of social programs.

5.0 METHODOLOGY

The topic of this chapter is the methodology to be adapted in this study. This study primarily aims to evaluate state R&D tax credit programs and therefore, it is important to demonstrate how to make casual inference of the program and its effects and how to make it more secure. Because state R&D tax credit programs can only be evaluated by using non-randomly selected samples that are non-randomly treated as a kind of social programs, the quasi-experimental design is the most suitable approach. The quasi-experimental design can be conducted through recognizing rival hypotheses based on the theory of validity, deciding the appropriate design from many possible quasi-experimental designs, and selecting the statistical methods depending on the selected specific quasi-experimental design.

In detail, the theory of validity provides the basis of judging the reliability of casual inference. The theory of validity primarily consists of (1) internal validity for reinforcing causality and (2) external validity for ensuring generalization, which are originally developed by Campbell (1957), and Campbell and Stanley (1963). Then, this theory is elaborated by adding (3) statistical conclusion validity and (4) construct validity by Cook and Campbell (1979) and Shadish, Cook and Campbell (2003). This theory also enumerates possible threats to validity, which are specific reasons indicating the possibilities of making wrong inference, under each type of validity. Therefore, the theory of validity is useful for understanding possible threats to validity for recognizing rival hypotheses in each specific case.

Next, based on this validity structure, two major social experiments for program evaluation, which are randomized experimentation and quasi-experimentation, are compared in terms of their strengths and weaknesses for dealing with threats to validity. Then, under the quasi-experimental approach, two possible specific designs, which are the non-equivalent control group design and the interrupted time-series design, are presented. The possible variations and combinations of these two types of designs are demonstrated as well.

Third, the specific econometric and statistical methods to be applied based on the above rationales are explored to find reliable and statistical empirical evidences. These methods include an analysis of covariance (ARCOVA) and an analysis of autoregressive integrated moving average (ARIMA) as simple methods and the difference-in-differences method and the matching method as elaborated ones.

After reviewing possible methods and the detailed approaches, quasi-experimentation is selected as an appropriate method in social setting and among many possible quasi-experimental designs, "the adjusted form of the interrupted time-series design by adding nonequivalent and untreated control group time series and switching replications," as the combined design of the non-equivalent control group design and the interrupted time series design, is chosen as the most appropriate approach for this study. For applying this selected quasi-experimental design, the difference-in-differences method and the matching method are chosen for this study.

5.1 MAKING CAUSAL INFERENCE

In order to make a definition of causal inference, it is useful to understand the logic of causation constructed by John Stuart Mill (Shadish, Cook and Campbell, 2003). Under his logic of causation, the following three characteristics should be satisfied: 1) the cause occurs before the effect follows, 2) the cause and the effect are relevant each other, and 3) there is no plausible other explanation for the effect other than the cause (Shadish, Cook and Campbell, 2003). In this sense, "causation" is different from "correlation" which corresponds to the second condition because simple correlations do not satisfy the other two conditions, which are the temporal precedence which corresponds to the first condition and the possible confounds which corresponds to the third condition. Indeed, when we do the scientific experiments for making causal inference, the cause should be manipulable. For illuminating causal inference, satisfying the above third condition, in other words, ruling out other plausible explanations is essential. Other "plausible" explanations are more important and more relevant explanations and therefore should be focused on in priority. In this sense, they should be distinguished from other "possible" explanations that are naturally endless.

The theory of validity is useful for identifying and recognizing plausible explanations, and broadly for making an inference. Validity is referred to "the approximate truth of an inference" or in other words "the extent to which relevant evidence supports that inference as being true or correct" (Shadish, Cook and Campbell, 2003). Importantly, the degree of validity varies and only can be judged approximately and tentatively due to the limitation of human knowledge and ability. There are two major types of validity, which are internal validity and external validity, originally introduced by Campbell (1957) and Campbell and Stanley (1963). Internal validity indicates the reliability of causal inference and external validity indicates the credibility of generalization of the revealed causal relationship. Therefore, these two kinds of validity have a conflicting relationship, which means an increase in the level of one validity results in a decrease in the level of the other validity.

Cook and Campbell (1979) and Shadish, Cook and Campbell (2003) also elaborated this structure into four different types of validity by adding statistical conclusion validity and construct validity.[62] Statistical conclusion validity indicates the reliability of statistical analysis for making inference and construct validity indicates the representativeness of the revealed causal relationship and the appropriateness of its general constructs. In this sense, the theory of validity provides the useful tools for judging and assessing the reliability of inferences including causal inferences, their generalizations and their conceptualizations into a higher order. Among four types of validity, internal validity is closely related to causation with a narrow definition. The four types of validity are presented in Table 17.

Table 17. Four types of validity

Types of validity	Definition
Internal validity	"The validity of inferences about whether observed covariation between A (the presumed treatment) and B (the presumed outcome) reflects a causal relationship from A to B as those variables were manipulated or measured" (p. 38)
External validity	"The validity of inferences about whether the cause-effect relationship holds over variation in persons, settings, treatment variables, and measurement variables" (p. 38)
Statistical conclusion Validity	"The validity of inferences about appropriate use of statistics for the correlation (covariation) between treatment and outcome" (p. 38)
Construct validity	"The validity of inferences about the higher order constructs that represent sampling particulars" (p. 38)

Source: Shadish, Cook and Campbell (2003, p. 38)

[62] The meanings of internal validity and statistical conclusion validity are the same as four major literatures mentioned above. However, the meanings of external validity and construct validity are little different in Cook and Campbell (1979) and Shadish, Cook and Campbell (2003). In Cook and Campbell (1979), external validity is limited to generalization of causal relationship to and across populations of persons and settings, and construct validity is limited to the representativeness of treatments and observations. In Shadish, Cook, and Campbell (2003), external validity is extended to generalization of causal relationship to treatments and observations, and construct validity is extended to the representativeness of persons and settings. For understanding distinctive features in detail, refer to Shadish, Cook, and Campbell (2003, p.38-39).

Under each type of validity there are threats to validity, which means specific reasons identifying the possibilities of weakening the logic of inferences and making wrong inferences. Threats to validity play a useful role in recognizing plausible and alternative explanations within an experiment in advance and subsequently building an appropriate experimental design to rule out them. Importantly threats to validity cannot be applied to every case in the same way. Rather, relevant and significant threats to validity vary in each case. Therefore, ruling out plausible threats to validity depends on identifying relevant threats to validity and identifying them depends on available knowledge and methodologies as well as specific contexts and backgrounds involved. In this sense, scientific experiments for drawing valid inference are featured by being quite qualitative rather than quantitative. The detailed lists of possible threats to each type of validity are presented below.

First, for strengthening statistical conclusion validity, it is essential to verify statistically the covariance of a presumed cause (A) and a presumed effect (B) *and* its effect size (i.e. level of significance or confidence intervals). Among three characteristics of Mill's logic of causation mentioned above, statistical conclusion validity is closely related to the second condition for reinforcing the correlation between A and B.[63] The threats to statistical conclusion validity mostly come from statistical methods to be applied. For example, the low level of significance cannot guarantee the existence of covariance. Basic assumptions of statistical methods to be used should be carefully examined. Measurement error is one of frequently observed factors to make a false conclusion. The possible threats to statistical conclusion validity, the corresponding possible results, and the possible solutions are presented in Table 18.

[63] As mentioned above, the third condition is closely relevant to internal validity.

Table 18. Threats to statistical conclusion validity

Threats to validity	Definition	Impact	Solution
Low statistical power	An experiment with low level of significance	Incorrect conclusion of insignificant relationship between treatment and outcome	- Using statistical methods such as blocking and matching - Using larger sample sizes - Correcting covariates - Improving measurement - Using a within-participants design
Violated assumptions of statistical tests	Violations of statistical test assumptions	Overestimating or underestimating the size and significance of an effect	- Correcting them by using the advanced statistical methods
Fishing and the error rate problem	Repeated tests without correction of the number of tests	Inflating statistical significance	- Correcting the number of tests and comparing them
Unreliability of measures	Measurement error	Weakening the relationship between two variables, weakening or strengthening the relationship among more than three variables	- Increasing the number of measurements - Improving the quality of measurements - Using techniques that detecting and correcting error variance
Restriction of range	Reduced range on a variable	Weakening the relationship between it and another variable	- Pilot testing measures and selection procedures for detecting the important data range
Unreliability of treatment implementation	Treatment which is implemented only partially or inconsistently	Underestimating effects	- Standardizing
Extraneous variance in the experimental setting	Possible extraneous factors to make influence within experimental setting	Inflating error	- Controlling them, measuring them and including them into the statistical model
Heterogeneity of units	The more heterogeneous units on an outcome variable	The larger standard deviations on that variable	- Using them as blocking with homogenous units or as covariates - Sampling the more homogeneous units (but possibly reducing external validity and possibly restricting the range of data)
Inaccurate effect size estimation	Inadequate statistical methods	Systematical overestimating or underestimating the size of an effect	

Source: Cook and Campbell (1979) and Shadish, Cook and Campbell (2003, pp.45-53)

Second, for reinforcing internal validity, which means making causal inferences, the three characteristics of Mill's logic of causation, mentioned above, need to be satisfied. These are a temporal precedence of a cause, a correlation between a cause and an effect, and no other plausible explanations. Especially, it is important to recognize and rule out the possible third factors (C) that threaten and weaken a causal relationship between a presumed cause (A) and a presumed effect (B), which is relevant to the third condition mentioned above. The possible third factors include other events occurring concurrently, called "history," systematic differences in characteristics between experimental observations and others, called "selection." In addition, natural changes over time, called "maturation," a change of measurement, called "instrumentation," and a tendency of moving extreme scores to the less-extreme scores, called "regression," are possible other factors to evoke confusion with a true treatment effect. Several threats to internal validity also could be combined, called as "additive and interactive effects of threats to internal validity." These all possible threats to internal validity could make a causal relationship weaken and result in a wrong inference. Accordingly they should be properly recognized and be ruled out depending on specific settings and specific contexts. The detailed list of threats to internal validity, the corresponding possible impacts, and the possible solutions are presented in Table 19.

Table 19. Threats to internal validity

Threats to validity	Definition	Possible impacts	Possible solutions
Ambiguous temporal Precedence	Lack of clarity about which variable occurred first reciprocal causation	Confusion about the cause and the effect	- Using statistical methods such as path diagrams or endogenous variables
History	Events occurring concurrently with treatment	The another possible cause to the observed effect	- Selecting comparison observations as similar as possible - Controlling the recognizable and significant events involved
Maturation	Naturally occurring changes over Time	Confusion with a treatment effect	- Selecting observations having similar development stages and similar environments
Testing	Exposure to a test affects test scores	Confusion with a treatment effect	- Allowing a use of different tests - Using a Solomon Four Group Design (dividing groups with/without and pretest/posttest)
Regression	Tendency of extreme scores to move less extreme scores	Confusion with a treatment effect	- Ruling out the criteria of selecting samples based on extreme scores - Using more reliable measures - Creating a large group in extreme scores - Using multivariate function of several variables - Using average value over time
Instrumentation	The change of the nature of measurement over time	Confusion with a treatment effect	- Using same instruments during a whole experiment period - Retaining all instruments and calibrating them
Selection	Systematic differences over conditions in respondent characteristics	The another possible cause to the observed effect	- Selecting comparison observations as similar as possible - Controlling the systematic difference by recognizing significant events involved in together
Attrition (Mortality)	Loss of respondents to treatment or to measurement	Producing artificial effects if that loss is systematically correlated with conditions	
Additive and interactive effects of threats to internal validity	The existence of more than two threats to internal validity	Combination of the impact of a threat and that of another threat	

Source: Cook and Campbell (1979) and Shadish, Cook and Campbell (2003, pp.54-61)

Third, external validity involves making inferences in which “a causal relationship holds over variations in persons, settings, treatments, and outcomes” (Shadish, Cook, and Campbell. 2003, p.83). Therefore, for securing external validity, the possible threats to validity to be ruled out could be listed in terms of units, treatments, outcomes, and settings, which are presented in Table 20. For increasing the possibility of generalization, using a diverse and heterogeneous group for making causal inference is recommended. Therefore, extending external validity could have a negative effect on internal validity. However, the possibility of generalization and the usefulness of its application are as significant as the high explanatory power of a causal relationship. Therefore, depending on a purpose of research, this trade-off relationship needs to be assessed.

Table 20. Threats to external validity

Threats to validity	Definition
Interaction of the causal relationship with units	“An effect found with certain kinds of units might not hold over other units” (p.87)
Interaction of the causal relationship over treatment variations	“An effect found with one treatment variation might not hold over other variations of that treatment, that treatment combined with other treatments, and only part of that treatment” (p.87)
Interaction of the causal relationship with outcomes	“An effect found on one kind of outcome observation may not hold over other outcome observations” (p.87)
Interactions of the causal relationship with settings	“An effect found in one kind of setting may not hold over other kinds of settings” (p.87)
Context-Dependent mediation	“An explanatory mediator of a causal relationship in one context may not mediate in another context” (p.87)

Source: Cook and Campbell (1979) and Shadish, Cook and Campbell (2003, p.86-90)

Fourth, for dealing with construct validity, which is referred to the representativeness from the sampling particulars to the higher-order constructs, the most important issues to be involved in are how to choose samples and how to manipulate experiments. The possible threats to validity, which is mostly relevant to “the match between the operations and the constructs used to describe those operations" (Shadish, Cook and Campbell, 2003, p.72), are summarized in Table 21. By using single operation (i.e. measure, manipulation, setting, and unit), the constructs could be measured irrelevantly and underrepresented, and by using single method, that method could be included as a part of the construct. Therefore, using multiple operations and multiple methods is recommended. Some threats to construct validity, including experimenters' expectations and participants' reaction, rivalry, and demoralization, are caused by the perceptions and the subsequent reactions of participants and experimenters during experiments and these could be a part of treatment construct. For fostering construct validity, constructs should be developed as the following processes: 1) a clear explication of construct, 2) a thoughtful selection of samples, 3) an assessment of the match between samples and constructs, and 4) a revision of constructs (Shadish, Cook and Campbell, 2003, p.66).

Table 21. Threats to construct validity

Threats to validity	Explanation	Impacts
Inadequate explication of constructs	"Failure to adequately explicate a construct" (p.73)	"Incorrect inferences about relationship between operation and construct" (p.73)
Mono-operation bias	"Any one operationalization of a construct" (p.73)	"Under-representing the construct of interest, measuring irrelevant constructs, complicating inference" (p.73)
Mono-method bias	"Using the same methods for all operationalization" (p.73)	"Becoming a part of the construct" (p.73)
Confounding constructs with levels of constructs	"Inferences about the constructs that best represent study operations" (p.73)	"Failure to describe the limited levels of the construct" (p.73)
Reactivity to the experimental Situation	"Reflecting participant's perceptions on responses" (p.73)	"Becoming a part of the treatment construct" (p.73)
Experimenter expectations	"Influence of the experimenter in participant responses" (p.73)	"Becoming a part of the treatment construct" (p.73)
Compensatory rivalry	"Motivating participants not receiving treatment to show their ability" (p.73)	"Becoming a part of the treatment construct" (p.73)
Resentful demoralization	"Motivating participants not receiving treatment to respond more negatively" (p.73)	"Becoming a part of the treatment construct" (p.73)

Source: Cook and Campbell (1979) and Shadish, Cook and Campbell (2003, pp. 72-81)

In sum, in order to make causal inference, which means that a presumed and preceding cause and a presumed and subsequent effect are correlated and other possible explanations are ruled out, the theory of validity is useful to assess this relationship. The theory of validity provides four kinds of validity to justify causal inference. These are internal validity, external validity, construct validity and statistical conclusion validity. In particular, internal validity for making a strong causal relationship and external validity for generalizing the revealed causal relationship have a trade-off and it is impossible to increase both of them at the same time. Therefore, the researchers should judge how to deal with this trade-off relationship. Under each validity, threats to validity, in other words, possible other explanations to weaken causal inference, are enumerated. These threats to validity are only be able to apply to each specific case in the way of specifically defining and modifying them depending on each specific setting and context.

In the next section, I examine possible approaches to make causal inference in social setting based on the theory of validity.

5.2 QUASI-EXPERIMENTATION

In this section, based on the theory of validity mentioned above, I examine the possible methods of social program evaluation. These are randomized social experimentation and quasi-experimentation. Based on the practical applicability, quasi-experimentation is the preferred method for evaluating social programs. Therefore, the detailed discussion of possible variations of quasi-experimental designs follows and among them, the most appropriate design for this study is selected.

5.2.1 Randomized social experimentation *vs.* quasi-experimentation

As mentioned above, the fundamental difficulty of social program evaluation, as one of the processes in making causal inference, stems from the fact that we cannot observe what would have happened without a program, in other words, the counterfactual.[64] To tackle this difficulty, one possible approach is randomized social experimentation, which means randomly assigning a treatment from a randomly collected sample by controlling social settings.[65] This randomness ensures a causal relationship between the outcome and the program, especially making alternative explanations for causal relationships implausible, and consequently producing a strong internal validity. In detail, the strength of randomization is the ability of avoiding selection bias which is one of threats to internal validity and the ability of avoiding the possible combination of a treatment effect and other possible threats to internal validity. Importantly, the randomness cannot rule out a possibility of maturation, history or regression, but it can reduce to confounding of these possible threats to validity with a treatment effect (Shadish, Cook and Campbell, 2003). However, there are several drawbacks to this approach. For example, controlling social settings and keeping apart a treatment group and a control group for avoiding

---

[64] If we can conduct one experiment with a program and at the same time, also conduct the other experiment without the program with controlling everything same, we can observe the effect of program easily by comparing the difference of outcomes. But it is impossible in the social world. What we observe is one having the program and the other not having the program but different from the first one. In other words, there are no observations having the program and at the same time not having the program, which is regarded as a missing-data problem (Blundell and Costa Dias, 2000). Therefore, even though we observe the increases of R&D spending after utilizing a program, we cannot know where this result comes from. The firms might increase its R&D spending without program in the economic expansion period or its own prompt necessity even in the recession without tax credit. One survey indicated the firms intended to raise R&D spending by 17 percent in 1982 despite the economic recession (McGraw-Hill, 1982; quoted by Bozeman and Link, 1984).

[65] There are two empirical analyses by using randomized social experiments. LaLonde (1986) did an empirical analysis of the effects of the employment program and Benus, Grover, Johnson, Shen and Wood (1995) did an empirical analysis of the Unemployment Insurance (UI) Self-employment Demonstration.

possible interaction, especially for a relatively long time, is extremely difficult. It is also extremely difficult to have a control group which is perfectly unaffected by any possible effects including indirect effects. Randomized social experiment is not applicable to extrapolation (Blundell and Costa Dias, 2000). It is also questionable to apply the result from this experimentation to a real social setting, which implies it has a weak external validity.

From the above reasons, quasi-experimentation which means evaluating programs without randomness, is a preferred, and in some senses a possible method in social setting. Quasi-experiments are defined as "experiments that lack random assignment of units to conditions but that otherwise have similar purposes and structural attributes to randomized experiments" (Shadish, Cook and Campbell, 2003). Although quasi-experiments are practically preferred and widely applicable in social setting, its nature of non-randomness evokes many plausible explanations and consequently makes it difficult to conclude a strong causality.[66] In other words, rival hypotheses, which are other plausible explanations for causal relationships with the presumed cause and the presumed outcomes, should be eliminated. To develop appropriate methods for dealing with this fundamental weakness is a major task of the quasi-experimental designs. In sum, program evaluation methods, which are one type of drawing a causal inference, are broadly categorized by randomized social experiments and quasi-experiments as following Table 22.

[66] The particular type of threats to internal validity, which is relevant to quasi-experimentation, is explored in the next section with an assessment of the possible quasi-experimental designs.

Table 22. The alternative program evaluation methods

Social Experiments	Randomized social experiments	Quasi-experiments
Strength	- Strong internal validity	- Practically preferred - Widely applicable
Weakness	- Hard to implement in social setting - Weak external validity	- Weak internal validity without ruling out plausible rival hypotheses - Essential to make alternative explanations of causal relationships implausible

Source: Campbell and Stanley (1963) and Shadish, Cook and Campbell (2003)

Accordingly, the quasi-experiments are more appropriate for evaluating social program than the randomized social experiments based on their practical applicability. However, there is a possibility of weakening causal inference in most quasi-experiments and therefore recognizing possible threats of validity is essential for successful experimentation. Indeed, these possible threats are varied depending upon the context in every experimentation. In next section, I examine how these possible threats are under the possible variations of the quasi-experiments and there are several possible approaches to make causal inference under the quasi-experimental designs.

5.2.2 Quasi-experimental designs

In this section, I discuss several possibilities of performing program evaluations following the quasi-experimentation approach. Although categorizing and naming possible quasi-experimental designs are little different across literatures, these can be broadly categorized and named as the non-equivalent control group design and the interrupted time-series design. The former design means having a treatment group and a control group with both pretest and posttest observation(s) and the latter design means having a large number of observations for the same variable over time. There are other possible quasi-experimental designs that either lack a control group or lack pretest observations on the outcome, which are also called as pre-experimental designs because these types of quasi-experimental designs do not ensure a causal relationship due to a lack of ability of ruling out plausible explanations of causal relationships (Campbell and Stanley, 1963).[67]

5.2.2.1 The non-equivalent control group design

The non-equivalent control group design,[68] which is the most common quasi-experiment, is characterized by comparing a treatment group and a control group, which are not randomly selected and therefore not equivalent. The non-equivalent control group design is also characterized by having both pretest(s) and posttest(s) of the same units. The basic form of the non-equivalent control group design is depicted in Figure 3. In Figure 3, O_j represents each observation (if j>0, O_j represents a posttest observation and if j<0, O_j represents a pretest observation) and X is a treatment. NR means non-randomly selected and treated. In this basic form of the non-equivalent control group design, two non-randomly distributed groups (NR) have two observations, one of which is a pretest observation (O_{-1}) and the other of which is a posttest observation (O_1), and the first group is treated with X as a treatment group and the second group is not treated as a control group. For example, O indicates a yearly observation for the industry R&D spending by state, and X is the adoption of the state R&D tax credit law.

[67] For the details, refer to Shadish, Cook and Campbell (2003).

[68] Shadish, Cook and Campbell (2003) named this type of quasi-experimental design as “quasi-experimental designs that use both control groups and pretests” and more specifically “the untreated control group design with dependent pretest and posttest samples” (p. 136).

$NR \quad O_{-1} \;\; X \;\; O_1$: *treatment group*

\- - - - - - - - - - - - - - - - - - -

$NR \quad O_{-1} \quad\;\; O_1$: *control group*

(where NR means means non-randomly selected and treated, O_{-1} represents a pretest observation, O_1 represents a posttest observation, and X represents a treatment)

Figure 3. The diagram of the basic form of the non-equivalent control group design

The non-equivalent nature makes it impossible to make a causal relationship only by comparing a difference or differences between these two groups. The fundamental threat of internal validity involved in this design is selection bias due to non-randomness, and it is highly expected that selection bias is combined with other threats to internal validity. First, selection-maturation is expected, in that there is a possibility of different growth patterns. Second, selection-regression can occur when two comparison groups have different distributions of level of outcomes based on selection criteria. Third, selection-history is expected in that each group could experience different events except treatment, which means that other possible factors are involved in each group in a different way and may affect a treatment effect. These plausible threats to internal validity to be possibly involved in should be assessed and ruled out based on contexts and knowledge available in each case.

There are possible variations under this type of quasi-experimental design for strengthening causal inference. One of them is adding one more observations for pretests, called as “untreated control group design with dependent pretest and posttest samples using a double pretest” (Shadish, Cook and Campbell, 2003, p.137). As represented in Figure 4, the dotted box indicates the basic form of the non-equivalent control group design, and the second pretest observations (O_{-2}), which occurred at the same time to each group with same time delay from the first pretest observation, are added. By comparing the difference of pretest observations, selection-maturation as one possible threat to internal validity can be assessed.

$$
\begin{array}{lllll}
NR & O_{-2} & O_{-1} & X & O_{1} \\
\hdashline
NR & O_{-2} & O_{-1} & & O_{1}
\end{array}
$$

Figure 4. The diagram of the untreated control group design with dependent pretest and posttest samples using a double pretest

The second possible variation is making treatments to both groups at different time periods, called as "untreated control group design with dependent pretest and posttest samples using switching replications" (Shadish, Cook and Campbell, 2003, p.137). In this design, control group also is treated but at the different time period as represented in Figure 5. Comparing the treatment effect of the first group and the treatment effect of the second group, both internal validity and external validity could be strengthen. This type of design can be improved if switched replications can be imposed to more than two groups. However, for successful experimentation, the treatment should be imposed to each group in a same way and be properly removed from the initial group ensuring no continuous effects.

$$
\begin{array}{llllll}
NR & O_{-1} & X & O_{1} & & O_{2} \\
\hdashline
NR & O_{-1} & & O_{1} & X & O_{2}
\end{array}
$$

Figure 5. The diagram of the untreated control group design with dependent pretest and posttest samples using switching replications

The third possible variation is "untreated control group design with dependent pretest and posttest samples using reversed-treatment control group" (Shadish, Cook and Campbell, 2003, p.137). In this type of design, both of two groups are received the treatments but the resumed effects of these treatments are opposite, such as enrichment and impoverishment. In Figure 6, $X+$ represents the treatment assuming to generate an effect as one direction, while $X-$ represents the treatment assuming to generate an effect as the opposite direction. Accordingly, construct validity of the effect could be reinforced in the sense that treatment effects could be assessed in two opposite ways. However, this type of design tends to be avoided in many cases based on ethical and practical reasons, in that the reverse-treatment generates harmful effects.

$NR \quad O_{-1} \quad X+ \quad O_1$

- -

$NR \quad O_{-1} \quad X- \quad O_1$

Figure 6. The diagram of the untreated control group design with dependent pretest and posttest samples using reversed-treatment control group

Indeed, these possible variations can be combined as "untreated matched controls with multiple pretest and posttests," "nonequivalent dependent variables," and "removed and repeated treatments" (see, in detail, Shadish, Cook, and Campbell, 2003, pp.153-156).

The non-equivalent control group design utilizes both pretest and posttest observations for two groups, which are comparable but nonequivalent, in order to make causal inference. This type of design can be reinforced and improved through using double pretests, using switching replications and using reversed-treatment control group. These possible variations can be combined as well.

In the next section, I examine the other type of quasi-experimentation mostly popularly used along with the non-equivalent control group design, which is the interrupted time series design, in detail.

5.2.2.2 The interrupted time series design

The interrupted time-series design is another popular quasi-experimental method along with the non-equivalent control group design. In this design, the observations are constructed on a relatively large number of observations for the same variable consecutively over time and the treatment occurs at a specific point in this series of observations. As a basic form, the simple interrupted time series means one treatment group with a large number of observations including before and after a treatment, which can be conceptualized as Figure 7. In Figure 7, O_j represents each observation for the same variable (if j>0, O_j represents a posttest observation and if j<0, O_j represents a pretest observation) and X is a treatment. For example, the effect of government subsidy (represented as X) on employment growth can be assessed through a historical change of number of employees, including several observations before subsidy (represented as O_{-5} through O_{-1}) and other several observations after subsidy (represented as O_1 through O_5). Compared with the non-equivalent control group design, the large number of observations for pretest and posttest are required in this type of design.

$$O_{-5} \quad O_{-4} \quad O_{-3} \quad O_{-2} \quad O_{-1} \quad X \quad O_1 \quad O_2 \quad O_3 \quad O_4 \quad O_5$$

Figure 7. The diagram of the basic form of the interrupted time series design

The usual threats to internal validity involved in this design are history and instrumentation. Because the interrupted time series design has many observations over time, there are plenty of rooms for the occurrence of other events and the change of measurements. Attrition and regression are other possible threats to validity. As mentioned above, the plausible threats to internal validity depend on the context and knowledge in each individual design and therefore building appropriate strategies to deal with this imperfect condition is the most important task of the quasi-experimental design.

As like the non-equivalent control group design, the interrupted time series design could be improved in some ways. One of them is "adding a nonequivalent no-treatment control group time series", which is a combination of the non-equivalent control group and the interrupted time series design (Shadish, Cook and Campbell, 2003, p 182) as shown in Figure 8. The major strength of this type of design is that the possible threats of internal validity including history and maturation can be tested by untreated control group time series.

$$
\begin{array}{cccccccccccc}
NR & O_{-5} & O_{-4} & O_{-3} & O_{-2} & O_{-1} & X & O_{1} & O_{2} & O_{3} & O_{4} & O_{5} \\
\hdashline
NR & O_{-5} & O_{-4} & O_{-3} & O_{-2} & O_{-1} & & O_{1} & O_{2} & O_{3} & O_{4} & O_{5}
\end{array}
$$

Figure 8. The diagram of the interrupted time series design adding a nonequivalent no-treatment control group time series

The second possible design is "adding nonequivalent dependent variables" and the nonequivalent dependent variable is defined as "a dependent variable that the treatment should not affect but that would respond in the same way as the primary dependent variable to a pertinent validity threat" (Shadish, Cook and Campbell, 2003, p. 182). Importantly, the second dependent variables are not only nonequivalent to the primary ones, but also different with some comparable aspects. In Figure 9, O_{ij} represents each observation for the same variable (if j>0, O_{ij} represents a posttest observation and if j<0, O_{ij} represents a pretest observation, and if i=a and i'=b, O_{ij} and $O_{i'j}$ are nonequivalent but relevant with each other) and X is a treatment. By collecting nonequivalent dependent variables, that are related to the primary dependent variable as changing and moving over time in a similar way to the primary dependent variable, but are not be affected by the treatment, the most threats to internal validity can be tested. The examples are daytime observations and nighttime observations, or weekday observations and weekend observations.

$$\begin{array}{cccccccccccc} NR & O_{a\text{-}5} & O_{a\text{-}4} & O_{a\text{-}3} & O_{a\text{-}2} & O_{a\text{-}1} & X & O_{a1} & O_{a2} & O_{a3} & O_{a4} & O_{a5} \\ \hdashline NR & O_{b\text{-}5} & O_{b\text{-}4} & O_{b\text{-}3} & O_{b\text{-}2} & O_{b\text{-}1} & X & O_{b1} & O_{b2} & O_{b3} & O_{b4} & O_{b5} \end{array}$$

Figure 9. The diagram of the interrupted time series design adding nonequivalent dependent variables

The third possible design is "removing the treatment at a known time" (Shadish, Cook and Campbell, 2003, p. 188). In figure 10, X means a treatment and $\cancel{X}$ means the treatment is removed. By removing the treatment, this type of time series design can reduce the many threats of internal validity, such as history, selection, attrition, and instrumentation.

$$O_1 \quad O_2 \quad O_3 \quad O_4 \; X \; O_5 \quad O_6 \quad O_7 \quad O_8 \quad O_9 \; \cancel{X} \; O_{10} \quad O_{11} \quad O_{12} \quad O_{13}$$

Figure 10. The diagram of the interrupted time series design removing the treatment at a known time

The fourth possible design is “adding switching replications” (Shadish, Cook and Campbell, 2003, p. 192) as shown in Figure 11. This type of design includes the time-series data for two nonequivalent groups and the treatment is treated to each group at different time periods. Because the treatment is treated at a different time, it can be analyzed in two different times, populations, and settings. Accordingly, this design can reduce the most threats of internal validity as well as reinforce the external validity and the construct validity.

$$\begin{array}{l} NR \quad O_1 \quad O_2 \quad O_3 \quad O_4 \; X \; O_5 \quad O_6 \quad O_7 \quad O_8 \quad O_9 \quad O_{10} \quad O_{11} \quad O_{12} \quad O_{13} \\ \hdashline NR \quad O_1 \quad O_2 \quad O_3 \quad O_4 \quad O_5 \quad O_6 \quad O_7 \quad O_8 \quad O_9 \; X \; O_{10} \quad O_{11} \quad O_{12} \quad O_{13} \end{array}$$

Figure 11. The diagram of the interrupted time series design adding switching replications

These possible variations can increase the validity of causal inference by removing and reducing possible threats depending on the possible manipulations and possible observations. The interrupted time series design utilizes the large number of observations of same variables over time for making inference and can be reinforced and improved through adding non-equivalent control group time-series data, adding non-equivalent dependent variables, and removing the treatment and adding switching replications.

In next section, among the possible quasi-experimental designs, the appropriate form of quasi-experimental design for this study will be selected.

5.2.2.3 The selected design in this study

The choice of an appropriate form of design depends on the nature of question to be addressed, the nature of context and knowledge to be involved, and available data. Table 23 shows the alternative quasi-experimental designs with possible variations, and indicates the application of this study.

Among the several quasi-experimental techniques discussed above, this study adopts a combined method of three possible and relevant approaches. First, the non-equivalent control group comparison is appropriate for this study because we have observations for both states having tax credits and states not having them. Again, the every state has a different environment for deciding its private R&D spending and employment, a different history of private R&D spending growth and employment growth and a different utilization of state R&D tax credit programs. These differences mean that the experimental groups and the control groups in this study are non-equivalent and therefore, possible threats to validity should be carefully examined and be reduced as much as possible.

Second, the interrupted time series design is combined with the non-equivalent control group comparison, for detecting outcome changes before and after utilizing tax credits with a large number of observations over time. Shadish and Cook and Campbell (2003) categorized this type of combination as “the interrupted time series design adding a nonequivalent no-treatment control group time series” (p. 182) as shown in Table 23. Campbell and Stanley (1965) named this type of quasi-experimental design as “the multiple time-series design” (p.55) and they explained that this type of design is the most superior design in quasi-experimental designs for dealing with possible threats of internal validity.

Table 23. The alternative quasi-experimental designs

Two alternative quasi-experimental designs	Possible variations	Application of this study
The non-equivalent control group design	"Untreated control group design with dependent pretest and posttest samples" (p. 137) $NR \quad O_{-1} \quad X \quad O_1$ $NR \quad O_{-1} \quad \quad O_1$	O
	"using a double pretest" (p.137) $NR \quad O_{-2} \quad O_{-1} \quad X \quad O_1$ $NR \quad O_{-2} \quad O_{-1} \quad \quad O_1$	X
	"using switching replications" (p.137) $NR \quad O_{-1} \quad X \quad O_1 \quad \quad O_2$ $NR \quad O_{-1} \quad \quad O_1 \quad X \quad O_2$	O
	"using reversed –treatment control group" (p.137) $NR \quad O_{-1} \quad X+ \quad O_1$ $NR \quad O_{-1} \quad X- \quad O_1$	X
The interrupted time series design	one treatment group with a large number of observations including before and after a treatment $O_{-5} \quad O_{-4} \quad O_{-3} \quad O_{-2} \quad O_{-1} \quad X \quad O_1 \quad O_2 \quad O_3 \quad O_4 \quad O_5$	O
	"Adding a nonequivalent no-treatment control group time series" (p.182) $NR \quad O_{-5} \quad O_{-4} \quad O_{-3} \quad O_{-2} \quad O_{-1} \quad X \quad O_1 \quad O_2 \quad O_3 \quad O_4 \quad O_5$ $NR \quad O_{-5} \quad O_{-4} \quad O_{-3} \quad O_{-2} \quad O_{-1} \quad \quad O_1 \quad O_2 \quad O_3 \quad O_4 \quad O_5$	O
	"Adding nonequivalent dependent variables" (p.184) $NR \quad O_{a-5} \quad O_{a-4} \quad O_{a-3} \quad O_{a-2} \quad O_{a-1} \quad X \quad O_{a1} \quad O_{a2} \quad O_{a3} \quad O_{a4} \quad O_{a5}$ $NR \quad O_{b-5} \quad O_{b-4} \quad O_{b-3} \quad O_{b-2} \quad O_{b-1} \quad X \quad O_{b1} \quad O_{b2} \quad O_{b3} \quad O_{b4} \quad O_{b5}$	X
	"Removing the treatment at a known time" (p. 188) $O_1 \quad O_2 \quad O_3 \quad O_4 \quad X \quad O_5 \quad O_6 \quad O_7 \quad O_8 \quad O_9 \quad \cancel{X} \quad O_{10} \quad O_{11} \quad O_{12} \quad O_{13}$	X
	"Adding switching replications" (p.192) $NR \quad O_1 \quad O_2 \quad O_3 \quad O_4 \quad X \quad O_5 \quad O_6 \quad O_7 \quad O_8 \quad O_9 \quad \quad O_{10} \quad O_{11} \quad O_{12} \quad O_{13}$ $NR \quad O_1 \quad O_2 \quad O_3 \quad O_4 \quad \quad O_5 \quad O_6 \quad O_7 \quad O_8 \quad O_9 \quad X \quad O_{10} \quad O_{11} \quad O_{12} \quad O_{13}$	O

Source: Cook and Campbell (1979) and Shadish, Cook and Campbell (2003, pp. 137, 145-148, and 181-195)

Third, the enactments of credits were occurred at the different time periods across states, and therefore switching replications can be utilized. Switching replications can be facilitated to increase the validity for both possible quasi-experimental designs as shown in Table 23. This type of design can reinforce both internal and external validities because the outcome changes are examined under different time periods, populations and settings.

In sum, this study utilizes the adjusted form of the interrupted time-series design by adding nonequivalent and untreated control group time series and switching replications. This type of quasi-experimental design can be diagrammed in the following Figure 12. In Figure 12, NR means that each group is non-equivalent, first three groups represent treatment groups by switching replications, and last group represents a nonequivalent and untreated control group. In detail, the states not having state R&D tax credit program are included in the control group (the fourth group) and the states having tax credit program are included in the treatment groups (the first three groups) depending on their enactment years.

$$
\begin{array}{llllllllll}
NR & O_1 & O_2 & O_3 & O_4 & O_5\ X & O_6 & O_7 & O_8 \\
NR & O_1 & O_2 & O_3\ X & O_4 & O_5 & O_6 & O_7 & O_8 \\
NR & O_1 & O_2 & O_3 & O_4\ X & O_5 & O_6 & O_7 & O_8 \\
\hdashline
NR & O_1 & O_2 & O_3 & O_4 & O_5 & O_6 & O_7 & O_8
\end{array}
$$

Figure 12. The diagram of the adjusted form of the interrupted time-series design by adding nonequivalent no-treatment control group time series and switching replications

In the next section, I examine the possible statistical methods for empirically manipulating the quasi-experimental design discussed in this section. In particular, I discuss the selected statistical methods for this study, which are the difference-in-differences method and the matching method, in detail.

5.3 THE STATISTICAL METHODS FOR IMPLEMNTING THE QUASI-EXPERIMENTAL DESIGNS

The possible quasi-experimental designs mentioned above are explanatory and conceptual. For finding empirical evidences with actual data, statistical and econometric methods, which are encompassed in the concept and the mechanism involved in, are required. Table 24 illustrates the alternative statistical and econometric approaches to the corresponding quasi-experimental designs.[69]

Table 24. The alternative quasi-experimental designs

<table>
<tr><th rowspan="2">Quasi-experimental designs</th><th colspan="2">Statistical methods</th></tr>
<tr><th>Simple ones</th><th>Elaborated ones</th></tr>
<tr><td>The non-equivalent control group design</td><td>-ANCOVA</td><td rowspan="2">-The difference-in-differences method
-The DD with covariates
-The difference-in-difference-in-differences method
-The DDD with covariates
-The matching method
-The matching method incorporating the DD
-The instrumental variable method
-Regression discontinuity method</td></tr>
<tr><td>The interrupted time-series design</td><td>-ARIMA</td></tr>
</table>

Source: Cook and Campbell (1979), Angrist and Krueger (1998), Blundell and Costa Dias (2000) and Bartik (2002)

[69] For overall assessment of possible statistical and econometric approaches for evaluation methods, refer to Blundell and Costa Dias (2000), Heckman, Ichimura, Smith and Todd (1998) and Angrist and Krueger (1998), which illustrated well-developed evaluation methods in labor economic literatures, and Bartik (2002) and Reed and Rogers (2003) in regional economic development literatures.

The analysis of covariance (ANCOVA) and the autoregressive integrated moving average (ARIMA) models are simple methods for manipulating two alternative quasi-experimental designs respectively.[70] The ANCOVA estimates the treatment effect of the non-equivalent control group design by comparing the posttest values of a treatment group and a control group and adjusting the differences of pretest values of two groups as follows: first, by defining the dependent variable as the value of the posttest, and using dummy variable indicating treated or not (D=1 if treated and D=0 if non-treated), the ANCOVA separately estimates the average change of a control group and the average change of a treatment group, and second, by adding a covariate which is the value of pretest, the ANCOVA adjusts the differences of two groups in pretests.[71]

The ARIMA estimates the treatment effect of the interrupted time series design by using dummy variable indicating after treatment and before treatment (D=1 if after treatment and D=0 before treatment) with estimating the movements of time-series data of outcome variable (defined as Y_t). For estimating the movements of time-series data of outcome variable Y_t, three kinds of identifications are used in the ARIMA: (1) differencing(s) adjacent observations (for example, $(Y_t - Y_{t-1})$ as first differencing, $\{(Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})\}$ as second differencing), (2) finding relationship between current observation and previous observation(s) (for example, $Y_t = \Phi_1 Y_{t-1} + a_t$ or, $Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + a_t$), and finding relationship between current observations and

---

[70] For understanding the ANCOVA model in detail, see Reichardt (1979) and for understanding the ARIMA model in detail, see McCain and McCleary (1979).

[71] The ANCOVA is specified as follows (Reichardt, 1979):

$$Y_{ij} = \mu + \alpha_i + \beta (X_{ij} - \overline{X}) + \varepsilon_{ij}$$

(where Y_{ij} = the value of posttest, μ = the overall mean, α_i = the treatment effect, β = the coefficient of the linear relationship between Y and X, X_{ij} = the value of pretest, $\overline{X}$ = the overall mean of the pretest, and ε_{ij} = the error term) previous random shock(s) (for example, $Y_t = a_t + \Theta_1 a_{t\ t-1}$ or, $Y_t = a_t + \Theta_1 a_{t-1} + \Theta_2 a_{t-2}$) (McCain and McCleary, 1979).

In addition, elaborated statistical methods, which include the difference-in-differences method, the matching method, the instrumental variable method, and the regression discontinuity method, can be applied to manipulate any possible combination of quasi-experimental designs described under the section 5.2.2. [72] The difference-in-differences method estimates the treatment effect by comparing the average change in posttest(s) and pretest(s) of the treatment group and the change in posttest(s) and pretest(s) of the control group with an assumption of that the change of the control group could capture the possible outcome changes not from the treatment. The matching method estimates the treatment effect by finding matches from sufficient observable variables and comparing their differences with an assumption of that each matched pair has no systematic difference except the treatment. The instrumental variable method estimates the treatment effect by adding instrumental variable(s) which do not affect the outcomes but do affect the participation with an assumption of that by doing so the possible bias from non-random participation can be removed. The regression discontinuity method estimates the treatment effect by adjusting and smoothing the possible discontinuity of outcomes resulted from a known underlying variable.

In the next section, I discuss the difference-in-differences method and the matching method in more detail. In addition, I discuss the difference-in-difference-in-differences method and the matching method incorporating with the difference-in-differences method, as the more elaborate ones. These methods are chosen for this study to evaluate state R&D tax credits.

[72] For the details, see Angrist and Krueger (1998) and Blundell and Costa Dias (2000).

5.3.1 The difference-in-differences method

The difference-in-differences method (hereafter, the DD method) measures the treatment effect by comparing the average outcome changes of a treatment group before and after a program over those of a control group. Namely, the DD method provides the magnitude of the relative changes between an experimental group and a control group within the four-cell frameworks as the combination of "before/after and treatment/control" as shown in the following Table 25.[73] If the change of a control group before and after a program can successfully capture the effect of other relevant factors, which are assumed to be similar to those of a treatment group, we can estimate the pure effect of a program.

However, a social program is not randomly distributed; in other words, a treatment group and a control group are not equivalent. Therefore, the change of a control group might not be able to properly measure the effect from other relevant factors. Thus, for controlling the inner causality within the DD specification, constructing valid control group are essential. As Blundell and Costa Dias (2000) pointed out, "common time effects across groups and no composition changes within each group" are important assumptions for the DD method. Many literatures (for example, Angrist and Krueger, 1998 and Heckman, Ichimura, Smith and Todd, 1998) suggested the comparability of a treatment group and a control group is a prerequisite for successful

[73] DD method is statistically specified as follows (Klette, Møen and Griliches, 2000);

$$\hat{\beta}_{effect} = (\overline{Y}_{t_1}^{treat} - \overline{Y}_{t_0}^{treat}) - (\overline{Y}_{t_1}^{control} - \overline{Y}_{t_0}^{control}) = \Delta\overline{Y}^{treat} - \Delta\overline{Y}^{control}$$

(where β = the treatment effect, Y_{t1} = mean of outcomes after treatment, and Y_{t0} = mean of outcomes before treatment)

And the mostly accepted empirical regression form is the following equation (Bertrand, Duflo and Mullainathan, 2002, p.2);

$$Y_{ist} = A_s + B_t + cX_{ist} + \beta T_{st} + \varepsilon_{ist}$$

(where Y_{ist} = the measure of outcome, A_s = fixed effects for county (or firm), B_t = fixed effects for year, β = the coefficient of the linear relationship between Y and X, X_{ist} = covariates, β = the treatment effect, T_{st} = dummy for posttest treatment group, and ε_{ist} = the error term) program evaluation by using the DD method. The other way to control for the causality is adding covariates. Meyer (1995) explained this method could adjust observable differences between different groups, as the regression-adjusted DD method. Wooldridge (2003) pointed out that it could control the possible systematic differences of two comparison groups over time and moreover, it could increase the explanatory power and decrease the standard error of the DD estimator by reducing the error variance.[74] Then, by applying the fixed-effect panel analysis, possible outcome changes from time-specific effects and cross-sectional fixed-effects, in other words, possible effects from unobservable variables, could be eliminated as well. Indeed, this paper also estimates White-corrected standard errors for obtaining the heteroskadasticity-consistent standard errors (White, 1980).

Table 25. The formation of the difference-in-differences method

$$Y_{it} = \beta_1 + \beta_2 T + \beta_3 Post + \beta_4 T \cdot Post + \varepsilon_{it} \quad \text{-----------------------------------} \quad (1)$$

(where Y_{it} = the measure of outcome, T=1 if treated, T=0 if non-treated, $Post$=1 if after program, $Post$=0 if before program, and ε_{it} = the error term)

	After treatment	Before treatment	Difference
Treatment	$\beta_1 + \beta_2 + \beta_3 + \beta_4$ (T=1 and Post=1) "A"	$\beta_1 + \beta_2$ (T=1 and Post=0) "B"	$\beta_3 + \beta_4$ (first differencing) "A−B"
Control	$\beta_1 + \beta_3$ (T=0 and Post =1) "C"	β_1 (T=0 and Post=0) "D"	β_3 (first differencing) "C−D"
	Effect of treatment on treated		β_4 (second differencing) "(A−B)−(C−D)"

(where $\beta_3 + \beta_4$ = the outcome change of the treated, β_3 = the outcome change of the non-treated, in other words, the outcome change from the relevant factors (especially unobservable factors) except program, and β_4 = pure effect of the program, namely outcome change only from the program)

[74] For the further understanding, see Wooldridge (2003, p432~438).

Meyer and Sullivan (2004) adopted the difference-in-differences method for evaluating the 1996 TANF reforms in the United States and they find the slight improvements on living standards of single mothers comparing with single women without children and married couples with children. Card and Krueger (1994) evaluated the New Jersey state minimum wage law by comparing employment of fast-food restaurants from four chains in New Jersey with that of similar restaurants in western Pennsylvania. They found that the minimum wage increase in New Jersey didn't result in the negative effects which are presumably assumed. Card (1990) estimated the effect of immigration (large-scale immigration from Cuba to Miami in 1980) on employment of non-immigrants by comparing employment changes of Miami and other similar cities (Atlanta, Los Angeles, Houston and Tampa-St. Petersberg). He found there was no significant increase in unemployment rate of Black people in Miami, which tells us that the expected negative effect on employment, in particular for low educated and low incomed people, was not observed.

In addition, the standard errors in the DD method should be carefully estimated. Bertrand, Duflo and Mullainathan (2002) pointed out the high possibility of serial correlation in the DD method and therefore, this paper estimates standard errors by "allowing for an arbitrary covariance structure over time within each state (p. 3)" for obtaining the serial correlation-consistent standard errors. Indeed, this paper also estimates White-corrected standard errors for obtaining the heteroskadasticity-consistent standard errors (White, 1980).

5.3.2 The difference-in-difference-in-differences method

Further, the difference-in-difference-in-differences method (the DDD method) estimates the effect of a program more precisely by separating the more affected group and the less affected group within a treatment group and a control group. Meyer (1995) suggested this specification as "having treatments that are higher order interactions" and explains these higher order interactions could capture "main effects" and "lower level of interactions effects." The DDD method is constructed by adding one more differencing of the outcome changes of subgroups within a treatment group after differencing from those of the corresponding subgroups within a control group as shown in Table 26.[75] Here, subgroups are defined as a directly and primarily influential group, and an indirect and less-influential group. Within this specification, the outcome changes of less-influential group in a treatment group after differencing those of corresponding group in a control group captures possible changes from unobservable variables which are not relevant to the program within the treatment group. In this sense, the DDD method results in better estimates than the DD method because this method can capture relevant effects which are able to still remain in the DD model.

[75] The DDD method is statistically specified as follows:

$$\hat{\beta}_{effect} = \left\{(\overline{Y}_{t_1}^{GroupA:treat} - \overline{Y}_{t_0}^{GroupA:treat}) - (\overline{Y}_{t_1}^{GroupA:control} - \overline{Y}_{t_0}^{GroupA:control})\right\} - \left\{(\overline{Y}_{t_1}^{GroupB:treat} - \overline{Y}_{t_0}^{GroupB:treat}) - (\overline{Y}_{t_1}^{GroupB:control} - \overline{Y}_{t_0}^{GroupB:control})\right\} = \Delta\overline{Y}^{GroupA:treat} - \Delta\overline{Y}^{GroupB:treat}$$

(where β = the treatment effect, Y_{t1} = mean of outcomes after treatment, and Y_{t0} = mean of outcomes before treatment)

Table 26. The formation of the difference-in-difference-in-differences method

$$Y_{it} = \beta_1 + \beta_2 Group + \beta_3 T + \beta_4 Post + \beta_5 Group \cdot T + \beta_6 Group \cdot Post + \beta_7 T \cdot Post + \beta_8 Group \cdot T \cdot Post + \varepsilon_{it}$$

(where Y_{it} = the measure of outcome, *Group*=1 if influential subgroup, *Group*=0 if less-influential subgroup, *T*=1 if treated, *T*=0 if non-treated, *Post*=1 if after program, *Post*=0 if before program, and ε_{it} = the error term)

	After treatment	Before treatment	Difference	
Influential group within treatment	$\beta_1+\beta_2+\beta_3+\beta_4+\beta_5+\beta_6+\beta_7+\beta_8$ (Group=1, T=1, Post=1) "A"	$\beta_1+\beta_2+\beta_3+\beta_5$ (Group=1, T=1, Post=0) "B"	$\beta_4+\beta_6+\beta_7+\beta_8$ (first differencing) "A−B"	$\beta_7+\beta_8$ (second differencing) "(A−B)−(C−D)"
Influential group within control	$\beta_1+\beta_2+\beta_4+\beta_6$ (Group=1, T=0, Post=1) "C"	$\beta_1+\beta_2$ (Group=1, T=0, Post=0) "D"	$\beta_4+\beta_6$ (first differencing) "C−D"	
Less-influential group within treatment	$\beta_1+\beta_3+\beta_4+\beta_7$ (Group=0, T=1, Post=1) "E"	$\beta_1+\beta_3$ (Group=0, T=1, Post=0) "F"	$\beta_4+\beta_7$ (first differencing) "E−F"	β_7 (second differencing) "(E−F)−(G−H)"
Less-influential group within control	$\beta_1+\beta_4$ (Group=0, T=0, Post=1) "G"	β_1 (Group=0, T=0, Post=0) "H"	β_4 (first differencing) "G−H"	
	Effect of treatment on treated		β_8 (third differencing) " {(A−B)−(C−D)} − {(E−F)−(G−H)} "	

Where $\beta_4+\beta_6+\beta_7+\beta_8$ (A−B) = outcome change of influential group within treatment, $\beta_4+\beta_6$ (C−D) = outcome change of less-influential group within treatment, $\beta_3+\beta_7$ (E−F) = outcome change of influential group within control, β_3 (G−H) = outcome change of less-influential group within control, $\beta_7+\beta_8$ (A−B)−(C−D) = outcome change of interesting group except the outcome change of less-influential group within treatment, which means outcome change closely related to program, β_7 (E−F)−(G−H) = outcome change of influential group except the outcome change of less-influential group within control, which means comparable outcome change from the relevant factors (especially unobservable factors) except program and closely related to program, and β_8 {(A−B)−(C−D)} − {(E−F)−(G−H)} = pure effect of program, outcome change only from the program

Collins (2003) applied this DDD method for testing the effect of state-level anti-discrimination laws. He differentiated black and white as subgroups within states having this law and states not having this law and found that the fair employment laws adopted in the 1940s had the larger effect than those adopted in the 1950s and the black women received the relatively larger benefits than the black men. Gruber (1994) adopted this DDD method as constructing the comparison group by using the following three dummies – state (having/not having laws), time (before/after), and the women of certain ages/others – for assessing the mandated employer provision of employee benefits such as the comprehensive health insurance coverage for maternity and he found the little effect on the wage increase. Yelowitz (1995) estimated the effects of the welfare program reform by using the DDD estimates with the following three dummy variables – state (having/not having laws), time(before/after), and the women having certain ages of children/others for differencing the effects of the necessity of the childcare – and he found the significant increase of labor force participation.

5.3.3 The matching method

As another way to deal with the heterogeneity between an experimental group and a control group, the matching method finds out the most similar observation(s) within a control group to each individual observation within an experimental group, and by doing so, each matched pair is assumed to have the same characteristics only except the treatment. The basic idea of the matching method is that it is possible to adjust the bias from differences between an experimental group and a control group by constructing matches through observed characteristics which determine either the participation, or outcome changes, and then averaging outcome differences of these matches. For successful matching, the following two important assumptions are necessary. One is the "unconfoundedness assumption" which means an assignment to the treatment is independent of outcomes conditional on observed variables, and the other is the "overlap assumption" or the "identification assumption" which means an assignment of the treatment conditional on observed variables has a probability bounded away from zero and one.[76]

Compared with the DD method as one of regression estimations, the matching method, as one of nonparametric estimations, could avoid strict assumptions which are, for example, the linearity and the homogenous treatment effect. Another advantage of the matching method is the capability of incorporating other methods. Therefore, in this study, the matching method incorporating the DD method is selected for evaluating the state R&D tax credits. For doing this, the dependent variable itself is constructed as the average change of outcomes and captured the possible effects from unobservable variables.

For better matching, however, it is desirable that the matching is performed by comparing multiple characteristics in order to reduce possible differences as much as possible. Therefore, there are several suggestions for capturing the degree of similarity of continuous variables and for creating a single dimensional comparable measurement from multidimensional properties. One of them is using Mahalanobis metrics, in other words, the covariates method or the multivariate method, which creates the unique distance based on the correlations between variables.[77] The matching method by using Mahalanobis metrics can reduce the differences between two comparison groups in any linear function by assuming that the covariance of two

[76] For the details, see Abadie and Imbens (2002) and Heckman, Hidehiko, and Todd (1997).

[77] For further understanding the matching method by using Mahalanobis metrics, refer to Friedlander and Robins (1995).

comparison group's observable variables are same (Rosenbaum and Rubin, 1985).[78] Another one is using the propensity score metrics based on conditional probability of being treated based on the assumption that observed characteristics are related to the probability to receive treatment.[79] The propensity score can be generated from the binary choice model for the participation. For using the propensity score metrics, the "balance property" assumption is required, in other words, the average propensity scores of treatment group and control group within each interval (equally divided intervals of the propensity score) should be not significantly different, which implies the treatment occurs independently with observed characteristics. While using Mahalanobis metrics might create a matched pair that is quite different from insufficient sample datasets, using the propensity score metrics could generate a matched pair having quite different characteristics but having a similar propensity score. Therefore, there is no absolute superiority upon these two methods. More recently, these two methods are combined (see, for example, Angrist and Hahn, 1999 and Zhao, 2004).

Another important issue related to the matching method is the possibility of creating pairs within a given sample. Because the matching is performed based on multiple variables including continuous variables,[80] the exact matching is almost impossible even after generating single dimensional measurement by using either the Mahalanobis metrics or the propensity score metrics. For dealing with this issue, there are several suggestions. First, we can do the nearest

---

[78] The Mahalanobis distance is defined as follows:

$$M(X',X'')=(X'-X'')^T V^{-1}(X'-X'')$$

(where X' means the vector of observable variables of treatment group, X'' means the vector of observable variables of control group, V means the covariance matrix of X.)

[79] For the properties of propensity score, see Rubin (1977) and Rosenbaum and Rubin (1983). For the details of the propensity score matching method, see Dehejia and Wahba (1999).

[80] Continuous variables could take on any possible type of values, for example, 12.029384 and 10.034, while discrete variables only take on a limited type of values such as 1, 2, and 3.

neighbor matching method based on the one-to-one matching, which is finding pairs as one treated observation and one controlled observation, which have the most similar characteristics based on the Mahalanobis distance or the propensity score. Second, we can allow the creating matches with one treated observation and more than one controlled observations, called as multiple matching. There are some possibilities to find a set of multiple matches; (1) specifying the number of matches, (2) including every observation in the control group within a certain radius of the Mahalanobis distance or the propensity score, (3) kernel matching, in which each treated observation is matched with every controlled observation with weighted distances, and (4) stratification matching, which is creating intervals and averaging values of every observation within an interval (Heckman, Ichimura, Smith, and Todd, 1998).[81] In addition, better matching could result from allowing replacements, which means a single observation in a control group could be matched with more than one observations in an experimental group. In this study, the nearest neighbor matching method based on one-to-one matching and multiple matching based on specifying number of matches is used. Indeed, in the matching method by using the propensity score metrics, a caliper is used for excluding the matched pairs having quite different characteristics by determining a maximum level of difference in propensity scores.[82]

Empirically, Isserman and Merrifield (1987) used the matching method for assessing the impact of a power plant establishment on regional economic growth as a growth pole. They found weak evidence of the overall increase in economic outcomes within the near-by areas and the only partial increase within limited industries and residents. They argued that these partial effects could generate other trickle-down effects after a certain time period. They also pointed

[81] For empirical implementation by using the STATA program, see Becker and Ichino (2002) and Leuven and Sianesi (2003).
82 For the details, refer to Dehejia and Wahba (2002).

out that this quasi-experimental approach is valuable for evaluating spatial economic development policies. Rephann and Isserman (1994) applied the matching method for testing the impact of highways and they found the interstate counties in close proximity to urban areas received relatively larger benefits, compared to rural interstate and off-interstate counties. Isserman and Rephann (1995) adopted the matching method for evaluating the Appalachian Regional Commission which performed a comprehensive economic development programs and they found significant faster growth of income, earnings and population than the control groups. Greenbaum and Engberg (1998) adopted the matching method incorporating DD method for measuring the effects of enterprise zones in six states. They found that the overall impact is small but positive impact for new establishments and negative impact for existing establishments. Mueser, Troske and Gorislavsky (2003) performed the various evaluation methods, which will be applied in this study as well, for estimating the effects of job training program in Missouri and they found that the matching method by using the propensity score metrics is most efficient.

In this chapter, I discussed the methodology for evaluating the effectiveness of state R&D tax credits. The most appropriate research design to be applied is the quasi-experimental design because for evaluating state R&D tax credit programs as one of social programs, outcome differences between states having tax credits and states not having tax credits, which are non-randomly assigned and non-randomly selected, will be compared. For conducting the quasi-experimental design, the theory of validity provides a useful guideline for recognizing plausible explanations caused from the non-randomness based on four types of validity, which are internal validity, external validity, construct validity, and statistical conclusion validity. Then, I analyzed the possible variations of quasi-experimental designs. The selected design of this study is the adjusted form of the interrupted time-series design by adding nonequivalent no-treatment control group time series and switching replications. In this study, the selected quasi-experimental design is represented by comparing states having tax credit program with different enactment years and states not having state R&D tax credit program. Then, the appropriate statistical methods for the selected quasi-experimental design are suggested as the difference-in-differences (DD) method and the matching method. The DD method estimates the program effect by comparing the average change of each group assuming the average change of a control group indicates possible change without program. The matching method estimates the program effect by comparing outcome differences by finding a set of mostly close matches from each group based on observable characteristics and averaging the outcome difference of each matched pair. By adding one more control group as influential and less-influential groups, which is the difference-in-difference-in-differences (DDD) method, and by using DD estimator for the matching method, which is the matching method incorporating the DD method, the estimates can be more precise.

In the next chapter, I conduct the research design for evaluating the effectiveness of state R&D tax credit programs. The research design is conducted based on understanding the theoretical backgrounds of the importance of R&D for economic growth overall as well as at the regional level and government intervention to private R&D discussed in Chapter 3, practical utilizations of R&D tax credit programs and their previous evaluations discussed in Chapter 4, and applicable methodology discussed in this chapter.

6.0 RESEARCH DESIGN

In Chapter 4, I examined the historical utilizations of state R&D tax credits. By 2003 thirty-eight states adopted their own R&D tax credit programs and these programs differ somewhat. These differences are analyzed by the base of the credit, the operation year, the application range and the state tax structure. In Chapter 5, I discussed the evaluation methodology. The selected methodology in this study is "the adjusted form of the interrupted time-series design by adding nonequivalent no-treatment control group time series and switching replications" as one of quasi-experimental designs and it will be conducted by the difference-in-differences/difference-in-difference-in-differences methods and the matching method incorporating the difference-in-differences method.

In this chapter, I develop the detailed evaluation strategies based on the above analyses. First, the hypothesis to be examined is defined clearly as the primary and most important task. This study tests two hypotheses. The primary one is testing a positive program effect of state R&D tax credit programs and the secondary one is testing different program effects by firm size.

Second, in the quasi-experimental setting instead of randomized social experimentation, I enumerate plausible rival hypotheses under five relevant environments to private R&D. These five environments are the state tax environment, the business environment, the R&D environment, the policy environment, and the firm environment. Understanding plausible explanations and ruling out them as much as possible make the causation strong and ensure.

Third, the evaluation strategies are developed under defining dependent variables, defining the experimental group and the control group, and developing evaluation methods. The selected dependent variables are R&D spending and employment and the selected observational levels of analysis are the state and the firm. The selected methods are the difference-in-differences/difference-in-difference-in-differences method, and the matching methods incorporating the DD method. The covariates and matching variables are chosen in order to deal with rival hypotheses. Based on the selected methods, six types of empirical specifications are listed. Then, the datasets to be used are illustrated. Finally, the summary of the selected methods and the selected levels of analysis is made for clarifying.

6.1 THE HYPOTHESIS

The primary hypothesis to be examined in this study is whether there is a positive effect from state R&D tax credits on private R&D spending and employment generation. A positive effect means an increase of private R&D spending and employment from the program, controlling for other factors. For testing this hypothesis, in other words, separating outcome change into one from the program and the other from other factors, it could be restated and modified as whether *the outcome changes of states having tax credits before and after operating programs* are larger than *those of states not having tax credits*. In this setting, the difference of outcome changes between two groups indicates pure program effects because we assume the change of the latter group captures the possible outcome changes from other relevant factors except program with a well-developed research design. If we find the larger increases in outcomes of the first group comparing with the second group, we can conclude the program has a positive effect.

The secondary hypothesis to be tested here is whether the effects of the state R&D tax credits differ by firm size. The effects of R&D tax credits on R&D spending and employment for small firms have been receiving substantial significances to policy makers as well as researchers. For policy makers, encouraging R&D activities of small firms is one of fundamental tasks for sustainable economic developments especially at the regional level. For researchers, the linkage between R&D activities and firm size is one of ongoing interests and policy impact related with this linkage is another interesting topic. This hypothesis is tested by separately estimating policy effects by firm size as the same method mentioned above and then comparing them. In the next section, I explore the possible rival hypotheses.

6.2 RIVAL HYPOTHESES

For testing the above hypotheses, a simple comparison of economic outcomes of the treatment group and the control group cannot guarantee the causal relationship of program utilization and its effect, because the treatment group and the control group are quite different, namely not equivalent in the quasi-experimental setting. This feature is an embedded weakness of most quasi-experimental approaches for evaluating social programs. For dealing with this nonequivalent nature, significant factors for constructing plausible other explanations, in other words, rival hypotheses, are developed. Defining and enumerating rival hypotheses are the most important tasks of every quasi-experimental approach because quasi experimentation only can justify its hypothesis and make it stronger by ruling out possible rival hypotheses as much as possible. These factors can be recognized by listing rival hypotheses, which are other possible causal relationships for resulting in different growth in R&D spending or economic outcomes without state R&D tax credits.

As shown in Table 27, possible rival hypotheses can be listed as five environments related to the state R&D tax credits. Selected five environments which affect private R&D activities are different tax environments, different business environments, different R&D environments, different government policies, and different firm characteristics.

Table 27. The rival hypotheses

Five environments	Detailed features
Different state tax environment	-Corporate tax Having/not having, an apportionment method, tax rate -R&D tax credit Having/not having, tax rate, base of credit, targeted/general, enactment year
Different business Environment	-Human capital Skilled worker (High education) -Initial economic condition Unemployment rate, income, population -Dynamic economic condition Population change, income change, employment change -Market size / Potential investment capacity Population, value added, GSP, PI -Agglomeration economies Establishment density, labor pooling, industrial mix
Different R&D environment	-Possibility of spillovers -Competition -Level of R&D -R&D performance of universities and research institutions -Federal level R&D support -Knowledge stock
Different government policy environment	-Direct funding Federal R&D funding, SBIR program, STTR program -Indirect Tax incentives Federal level/state level -Other economic development strategies Enterprise zone, science park
Different firm Characteristic	-Sales -Firm size Number of employee -Establishment status

I examine rival hypotheses in detail in the next section. Briefly, these five environments, as plausible alternative explanations for different outcome changes before and after program across states, are introduced as follows.

First, the overall tax environment is closely related to the tax credit because the R&D tax credit is naturally interrelated with other tax provisions. In this study, the features particularly related to R&D activities are selected and examined. Each state has developed its own tax structure over time depending on demand and necessity and therefore, each state tax system is distinguishable in almost every aspect. This feature makes it difficult to compare them and especially the comparative analysis with time-dimension is quite complicated.[83]

Second, the business environment generates different business outcomes and therefore provides many rival hypotheses. For example, both employment changes and income changes result from the overall investment capacity, the overall economic condition, the level of agglomeration, and the level of human capital, which are quite varied depending on time and region. The selected factors are essential determinants for economic growth as well. Based on these casual relationships, possible measurements for these factors in this field will be included in the empirical specification in this study.

Third, the R&D environment, such as the level of R&D and the possibility of spillovers, provides fundamental conditions for major outcomes of the program that are examined. It could be a part of the above business environment, however this field is separately considered in this study. These business and R&D environments can be understood particularly well within the regional context. Therefore, finding the comparison group at the MSA level or at least at the county level is quite reasonable and also it could construct the better comparison group.

[83] For the details, see Hall and Wosinska (1999).

Fourth, government policies could affect economic outcome changes as well. In particular, the regional level efforts for economic development have been largely growing and the state R&D tax credit is one of them. We expect economic outcomes are affected by these policy environments in various ways. These policies aim to attract new firms and people within a region, create new job opportunities, and increase the level of income and output. Although it is hard to define economic development strategy itself and measure its accomplishment, R&D related strategies are selected and examined. At the federal level, SBIR funds and government direct funding are another factors that are influential in determining private R&D activities. Some features within this environment are related to first one, the state tax structure, however, the specific policies for economic developments are considered under this category.

At last, the firm characteristics that are particularly important in the firm-level analysis are sales, number of employees and firm size.

Next, each environment is examined in detail, with the perspective of constructing valid control groups and selecting measurable variables as covariates or matching variables. In addition, rival hypotheses, that are difficult to be ruled out in empirical specification, are also noted. For doing this, significant determinants of urban and regional economic growth, which are mentioned above, are included under each environment.

6.2.1 The different state tax structure

Different state tax structures might result in different outcome changes. Based on the assumption of private firms' sensitive reaction to the tax system, if state A doesn't impose corporate tax on specific expenses, no tax is better than a tax credit, which is only a partial deduction to corporate tax. In other words, states without corporate taxes could provide better environment for private economic activities, including R&D activities. Therefore, states that don't have corporate taxes, which are Nevada, South Dakota, and Wyoming, should be eliminated from control groups.

Among 48 states having corporate tax system, 38 states utilized state R&D tax credit programs by 2003 and these programs has been operated different somewhat. Their different features are analyzed in the section 4.3.1 in detail. The related issues are a base of tax credit, an enactment year, and an application range. These differences provide fundamental criteria for constructing the control group and the experimental group in this study.

Moreover, following Rashkin (2003), "an apportionment method" could affect private R&D activities as one of the most important features of state corporate tax systems relating to the state R&D tax credit. An apportionment method is a method indicating how to calculate corporate income tax from possible three factors, which are income, sales and property. If corporate income tax is calculated based on income rather than property, it could provide structural benefit for R&D activities, in the sense that initial R&D is more related to building new facilities or hiring new workers rather than increasing income. Therefore, compared to three-factor formula evenly using three factors, single-factor formula or double-weighted formula based on sales provides a structural incentive for R&D activities (Rashkin, 2003). Therefore, states only using similar apportionment methods could construct valid control group as well.

Besides the different rate of corporate tax, a different rate of R&D tax credit might also result in different economic outcome. According to the previous studies, 1% increase of the federal level R&D tax credit generally result in 1% increase of private R&D spending, although there are some variations. In reality, the rates of state corporate taxes as well as state R&D tax credits are varied by state and both of them have been modified several times. For example, Michigan imposes 1.9% of corporate tax and 6.5 % of R&D tax credit but California imposes 8.84 % and 15 % of corporate tax and R&D tax credit, respectively in 2004. If we even consider different modifications of these corporate tax rates and tax credit rates by state over time, it is almost impossible to control these factors within the specification of this model. Therefore, in this study, these issues can only be dealt with by the unexplained part or unobservable variables. By doing so, the sensitivity depending on tax rate cannot be provided. This is one of weakness of this study, in the sense that the sensitivity of tax credit rate is one of important information for the policy implementation. However, it could be also one of advantages, in the sense that we don't need to compare the different tax rate and the tax credit rate of each state and it makes quantitative evaluation more feasible.

6.2.2 The different business environment

Different business outcome growths could result from different business environments primarily. The argument within this field is closely related to the above arguments about fundamental factors of urban and regional economic growth. First, human capital is most fundamental ingredient of economic growth and in particular, the mobility of high educator is prerequisite of performing R&D.[84] Indeed, firms that encounter a lack of skilled workers within a region simply couldn't increase their manpower, which is known as a place-mismatch or a skill-mismatch. Therefore, human capital is one of the important R&D environments. Second,

[84] Human capital could be constituted by the overall labor force. However, in this study the term of human capital is used for defining the skilled worker particularly for R&D and innovation and population change and labor force is categorized under overall economic condition.

economic conditions, such as labor force, per capita income, population change, and unemployment rate, might have strong linkages with business outputs. As a dynamic condition, transition of these factors is also one of economic indicators capturing an economic trend of a region. Next, overall current investment capacity is another basic factor for determining future R&D investment and new job creation. Then, agglomeration economies are also recognized as one of most important impetus of regional economic growth. Regional deviations of overall R&D investment and employment depend on primarily regional deviations of these business factors, rather than government policy tools including R&D tax credits. Therefore, controlling the effects from these factors is important for estimating the accurate policy impacts. Based on the significance and data availability, these factors are included as either covariates or matching variables.

6.2.3 The different R&D environment

Different R&D environments might affect private R&D activities as well. Knowledge spillovers are particularly important for R&D activities, because they can generate multiplier effects by sharing knowledge without decreasing returns and yield the higher rate of social returns than the private rate of returns, and consequently provide the legitimization of government interventions. This feature could be better considered within a regional context because localized knowledge spillovers are generally accepted in theories and empirics (see, for example, Lucas, 1988; Jacobs, 1969; Jaffe, Trajtenberg and Henderson, 1993). Even though the communication and transportation systems have been dramatically developed, the face-to-face interactions and co-locations within the proximity are essential for gaining the benefits from knowledge spillovers. Especially Jaffe, Trajtenberg, and Henderson (1993) found the empirical evidence of the importance of proximity for knowledge spillover based on patent citations and this study tells us that the closeness is fundamental for create new patents from the previous patents, which are gathered and filed with highly well-organized system and therefore easily accessible anywhere.

Although spillover is one of important for R&D and economic growth, it is hard to be measured and detected (see, for example, Krugman, 1991b). Therefore several proxies are suggested for controlling possible difference of spillovers by region. One of these proxies is the establishment density, which also possibly captures the degree of agglomeration. Indeed, because the knowledge spillover could be intensified through the higher interaction, the degree of density, in other words competition, is one of important R&D environment factors. The degree of competition also can be measured by the establishment density or the population density as like the degree of agglomeration.

It is also generally believed that the level of R&D differs by region and this overall R&D capacity of region makes a difference of R&D investment and regional economic growth. The level of R&D could be measured by the knowledge stock by using patent data or R&D expenditure (see, for example, Crosby, 2000; Griliches, 1992; Jaffe, 1986). In other way, regions having a similar level of universities or research institutions could be considered as a valid control group. Several empirical studies such as Jaffe (1989) and Anselin, Varga and Acs (1997) found the evidence of the significant role of universities for innovation through knowledge spillover.

Moreover, skilled workers, in other words, human capital, which is recognized as one of important business environment, could be one of important R&D environments because they create the idea and through knowledge spillovers.

6.2.4 The different government policy environment

Government policies are also assumed to affect economic outcomes. Namely, various government interventions in private R&D activity and other overall regional economic development strategies could affect outcome changes. First, federal government intervenes in private R&D activities either by direct funding (i.e. federal R&D funding,[85] SBIR Program, and STTR program) or by indirect tax incentives (i.e. federal R&D tax credit). The SBIR program specially has targeted small companies for encouraging their R&D activities since 1983. Similar with SBIR program, STTR program also has provided special funds for small companies starting 1990. Depending on R&D capacity of each firm and region, the amount of each funds are quite varied as well by region (most of data is available at the state level). It is assumed that these funds from federal government make an influence to deciding the amount of a private company's own R&D funds or other economic outcomes (especially employment as one of dependent variables in this study). Indeed, these funds are one of the indicators of regions' R&D capacity and could generate multiplier effects combining with private R&D investments. Therefore, these variables are included in empirical specification as covariates or matching variables in terms of the amount of funds.[86]

Next, there are lots of government interventions at the local level. Since the 1970s, local governments have performed a variety of policy tools under the name of local economic development strategies. The local economic development policies have substantially modified in

[85] Following the categorization of NSF datasets, federal R&D funding is defined as the R&D expenditures financed by the federal government and performed by the private companies. Within this categorization, local government R&D funding, which is the R&D expenditures financed by the local government and performed by the private companies, is separately defined.

[86] STTR program is excluded, because it started in 1990 but the matching year is chosen before 1990, based on the theoretical suggestion which tells us a control group should be similar with an experimental group at least before program.

terms of their meanings and their applications. Broadly these efforts are categorized as following: 1) smokestack-chasing efforts for inducing firms within regions, 2) entrepreneurial efforts for encouraging innovation and technology-related industry, and 3) “third wave” efforts as playing a role of mediators (Eisenger, 1995). Examples of these “third wave” programs include science parks, enterprise zone, and supporting structures for high-technology industry. Because these efforts are interwoven and conducted together as a comprehensive and extensive way, and mostly accomplished for a long time period, examining each fragmented local economic development strategy is so time-consuming and especially analyzing them historically is quite difficult due to an unclear time line of each program. From this aspect, in this study, these overall local economic development strategies are dealt with indirectly, as unexplained parts.

6.2.5 The different firm environment

Firm characteristics might make differences in the firm-level R&D and the utilization pattern of the R&D tax credit. As one of economic actors, firms deliberately make a decision based on their capacity and internal economic conditions, which can be is distinguished with regional level economic conditions. In detail, the firm level characteristics, such as a firm’s sales and a firm’s number of employees, could be distinguished with the regional level measurements, such as GSP and regional labor force. There are also some arguments in which firm characteristics play an important role for utilizing R&D tax credits. For example, from anecdotal analyses (for example, Sigalla and Viard, 1999), the most beneficiary firms from the state R&D tax credits are large firms.[87] This argument will be tested by separately measuring the effect of tax credits by firm size. This is meaningful in the sense that the relationship between firm size and R&D activity is one of interesting issues in the R&D-related literatures along with the expected different effects of tax credits by firm size. Establishment-status also might affect the ability to claim the credit for the following reasons; qualified R&D expenditures for applying the tax credit are determined based on previous R&D expenditures, and tax credits are only able to be claimed from corporate taxes, which are only available from profitable outputs.[88]

6.2.6 Summary

Figure 13 summarizes important factors for deciding private R&D activities and employment, and their relationships with state R&D tax credit programs, based on the above rival hypotheses with listed relevant environments. In Figure 13, two different kinds of outcomes (i.e. R&D expenditure and employment) which will be examined as dependent variables in this study, and state R&D tax credit programs which are a major source of generating regional outcome differences, are highlighted.

Then, the relevant rival hypotheses, namely the five environments discussed above, are highlighted outside, indicating these environments are interwoven, consequently determining economic outcomes. Within each environment, several economic factors, which are important and measurable, are chosen for empirical specification.

---

[87] OTA (1985) provided the evidence of the similar skewed tendency of large firms' receiving the most of federal R&D tax credits.

[88] Refer to Hall and Wosinska (1999), for understanding different situations between existing firms and new firms for applying the state R&D tax credit.

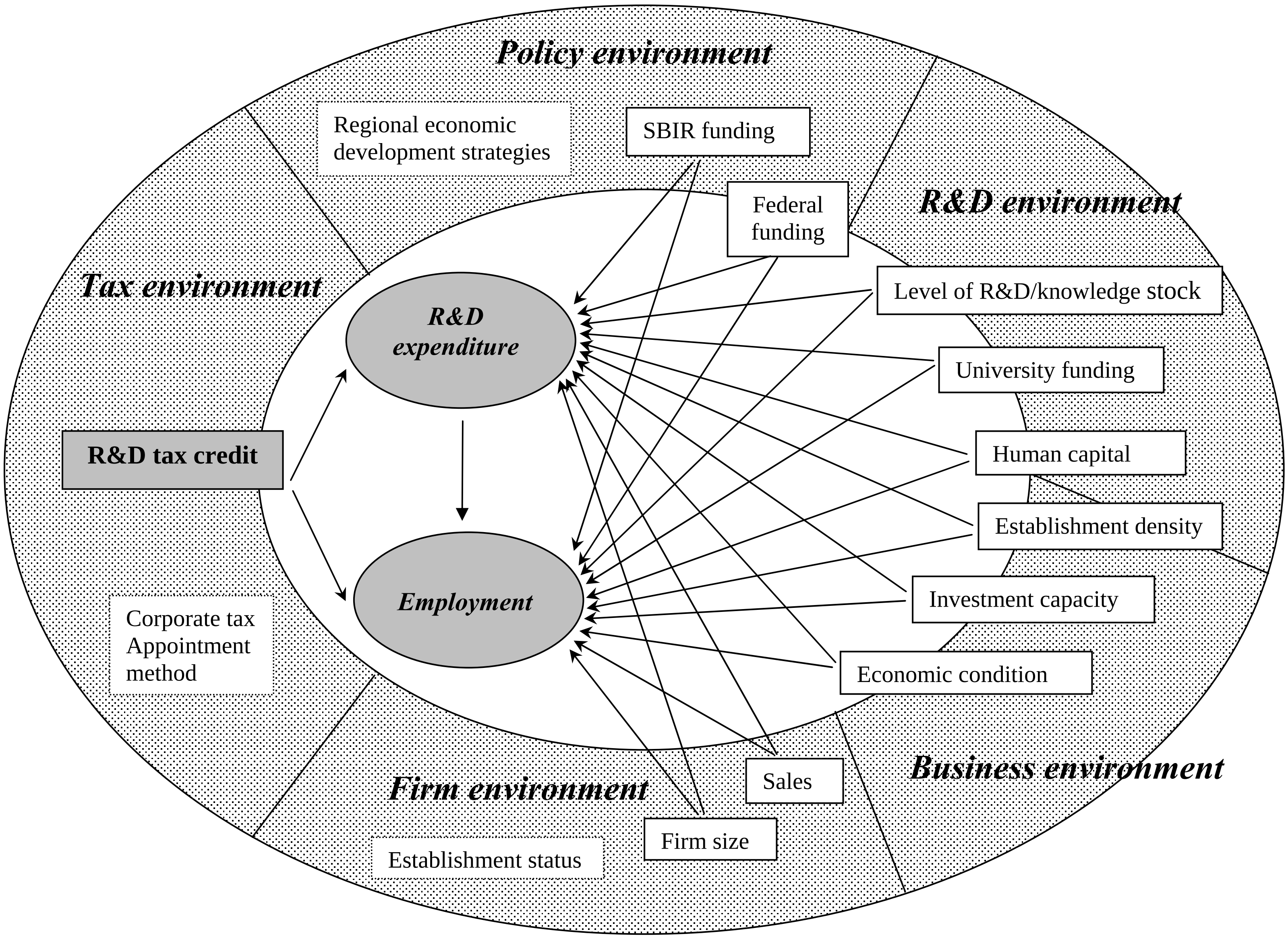


Figure 13. The relationships between rival hypotheses and program outcomes

These selected factors make it possible to rule out (at least deal with) corresponding rival hypotheses by adding covariates or matching variables within empirical specifications. These factors are indicated by a bold lined box with a line connected with corresponding outcome variables. However, there are some factors relevant to rival hypotheses that couldn't be measured directly and therefore only be ruled out as unexplained parts. These factors are indicated by a dotted box with no line connected with outcome variables.

It is assumed that all selected economic factors affect two outcomes and importantly selected outcomes have their own casual relationships (i.e. R&D expenditure → employment). Some factors, which are included in two environments, are also indicated. For example, human capital and establishment density are important factors in business environment and R&D environment simultaneously. In the next section, I develop the detailed evaluation strategies for dealing with rival hypotheses.

6.3 EVALUATION STRATEGIES

As the next step, the detailed evaluation strategies for eliminating the above rival hypotheses are enumerated. First, dependent variables are defined and the level of analysis including definitions of firm size and the high-technology industry is discussed. Second, the treatment group and the control group are defined based on the hypotheses to be tested. For better comparison some selection rules to make these two groups comparable are developed and adapted. In addition, an important assumption for comparing them and measuring the program effect is illuminated. Third, the selected methods, which are the DD/DDD method and the matching method incorporating the DD method, are developed in detail. The covariates and the matching variables are selected to deal with rival hypotheses. Then, the empirical specifications to be tested are enumerated and the data sources to be used are listed. Finally, the multiple analytical methods and estimators are summarized before the analyses are described.

6.3.1 Defining dependent variables

The expected economic outcomes as dependent variables with data sources are shown in Table 28. The changes before and after the program are measured by two outcomes: 1) the changes in R&D spending and 2) the changes in level of employment. The latter economic outcome could not result solely from state R&D tax credits, however, this study includes the analysis of effects on employment levels in order to ascertain the overall effects of state R&D tax credits in the grounds of employment growth, which is a primary concern of nearly every regional economic development strategy. For measuring this effect, this study does utilize various evaluation strategies which allow separate the main effect from the possible effects from other relevant factors and indeed it is additionally assumed that the amount of private R&D expenditure also affects the level of employment as depicted at the above Figure 13.

Table 28. The expected outcomes with available datasets

Expected Outcomes	Data sources	
	Firm	State
R&D spending	- Compustat by Standard & Poor's	- Total funds for industrial R&D performance by the NSF
Employment	- Compustat	- Current Employment Statistics (CES) by the Bureau of Labor Statistics (BLS)

These two possible outcomes (i.e. R&D expenditure and employment) are constructed either at the state level or at the firm level. The firm level analysis could control the different firm characteristics as a rival hypothesis under the section 6.2.5 and makes it possible to test the different effect by firm size which is the secondary hypothesis in this study. Within the firm-level analysis, the firm size is defined based on the number of employees for testing whether there is any difference of program effects by firm size. Small, medium and large firms are firms having less than 200 employees, between 200 and 500 employees and over 500 employees respectively. We use the base year for matching variables as 1990 and therefore the number of employees in 1990 is the criteria for deciding the firm size.

These variables are also broken down at the 3-digit level of Standard Industrial Classification (SIC) system except the state-level R&D spending data.[89] The industry-level analysis makes it possible to divide the high-technology industry group and the non-high-technology industry group and therefore to construct the valid treatment group and the control group in the sense that the high-technology industry is more relevant to the R&D tax credits. It also makes it possible to perform the DDD methods. By doing this, the better estimates are provided in the sense that multiple estimates can provide multiple evidences and also can complement each other by dealing with the different possible rival hypotheses.

An important issue for using the industry-level data and dividing the high-technology industry and non-high-technology industry is that there is no clear definition of the high-technology industry, even though this terminology has been frequently mentioned and specified (for example, Wallsten, 2001; DeVol, 1999; Lugar and Goldstein, 1991). It used to be broadly defined as “industry engaged in technically sophisticated activities that lead to product or process innovations, new inventions, or more generally, the creation of knowledge” (Goldstein and Luger, 1997) and categorized on the basis of either R&D expenditures or number of scientists

[89] The state-level R&D spending data is provided by the NSF and is only available as the aggregate level.

and engineers. In this research, the high-technology industry is classified by occupational criteria as following the Bureau of Labor Statistics (BLS) work, in which "R&D-intensive industries have at least 50 percent higher number of R&D workers than the average proportion for all industries" based on the BLS Occupational Employment Statistics surveys from 1987 to 1989 (Hadlock, Hecker and Gannon, 1991).[90] Here, the industry is defined at the 3-digit SIC level,[91] and R&D workers are defined as workers who spend most of their time for R&D activities. Among these industries, the defense-dependent industries are excluded,[92] because this study focuses on private R&D. From the above previous BLS work, the selected high-technology industry is listed in Table 29.[93] Then, the other remaining industries in manufacturing industry group (i.e. SIC 20~39) and service group (i.e. SIC 70~89) at the 3-digit SIC level are categorized by less R&D-intensive industries.

---

[90] Before this classification, the BLS also made the earlier version.(1983), in which R&D employment and R&D expenditures were used as criteria.

[91] Luker and Lyons (1997) pointed out the difficulty of classifying the high-technology industry at the industry level because there is a huge heterogeneity of individual firms within an industry.

[92] From the following BLS work by Luker and Lyons (1997) "defense dependent industries are defined as at least 50 percent of the output is for defense in 1987, which is the most recent peak year for defense expenditures." Those are aircraft parts (SIC 372), guided missiles and space vehicles (SIC 376) and search and navigation equipment (SIC 381).

[93] These selected industries are mostly same as those in other relevant studies such as Browne (1983), Malecki (1984) and Riche, Hecker and Burgan (1983) (Luker and Lyons, 1997).

Table 29. The list of the high-technology industry

SIC	Definition	SIC	Definition
131	Crude petroleum and natural gas operations	355	Special-industry machinery
		357	Computer and office equipment
211	Cigarettes	362	Electrical industrial apparatus
281	Industrial inorganic chemicals	366	Communications equipment
282	Plastics materials and synthetics	367	Electronic components and accessories
283	Drugs	371	Motor vehicles and equipment
284	Soap, cleaners, and toilet goods	382	Measuring and controlling devices
285	Paints and allied products	384	Medical instruments and supplies
286	Industrial organic chemicals	386	Photographic equipment and supplies
287	Agricultural chemicals	737	Computer & data-processing services
289	Miscellaneous chemical products	871	Engineering and architectural services
291	Petroleum refining	873	Research and testing services
335	Nonferrous rolling and drawing	874	Management and public relations services

Notes: Among the originally selected 30 industries, three of them (i.e. aircraft and parts (SIC 372), guided missiles and space vehicles (SIC 376), and search and navigation equipment (SIC 381)) are excluded because these industries are defense-dependent and two of them (i.e. miscellaneous petroleum and coal products (SIC 299) and services, not elsewhere classified (SIC899)) are excluded because the publishable data is not available.
Source: Hadlock, Hecker and Gannon (1991), Luker and Lyons (1997)

The industries are classified by the Standard Industrial Classification (SIC) system instead of the North American Industry Classification System (NAICS) in this study. The SIC system is more useful for analyzing historical data because the SIC system has been widely used since the 1930s with the latest revision in 1987, until the NAICS, which is the newly developed industrial classification system, was introduced in 1997 and has replaced the SIC system.

In addition, two possible outcome variables are measured by two values, which are level and log. By doing so, program effects can be measured by 1) the increase of amount of R&D spending or number of employees from level analysis and 2) the % increase of R&D spending or number of employees from log analysis.

6.3.2 Defining the treatment group and the control group

The treatment group and the control group are defined based on the hypothesis to be tested, which is whether the outcome changes of the treatment group before and after the state R&D tax credit programs are larger than those of the control group. The treatment group is defined as states having tax credits, and the control group is defined as states not having them. In the case of the firm level analysis, the treatment group constitutes firms within states having tax credits, and the control group constitutes firms within states not having credits. One important issue regarding defining the treatment group and the control group is that states have quite different tax systems as discussed in Chapter 4, section 4.3. This issue is also relevant to ruling out the possible rival hypotheses from the different tax environment.

Based on the notion that the treatment group and the control group should be comparable, only states confirming next three conditions will be classified into the treatment group; (1) states using the state tax system of "single-factor formula or double-weighted formula based on sales", (2) states having the R&D expense-based credits and allowing credits to every qualified R&D spending, and (3) states enacting credit laws al least until 1999. Then, only states confirming next three conditions will be categorized as the control group; (1) states having the same state tax system as above, (2) states not having tax credits *or* states having the R&D expense-based credits but allowing credits to only limited R&D spending based on the industry and area, and (3) states enacting credit laws after 1999. Other states not being satisfying the above criteria are excluded in this study.

The year of 1999, which is a base year for separating the treatment group and the control group, is chosen based on data availability and more importantly for the purpose of constructing the better comparison group. Namely, by doing so, among the top ten states performing R&D

activity in US, six of them are constituted in the experimental group and the rest of them are consisted of the control group, which provides one of the fundamental bases for an appropriate comparison. In detail, some states performing the significant portion of private R&D activities in the US but enacting the credit law after 1999 (i.e. Michigan, Ohio and Texas) are categorized as the control group and consequently two comparison groups are more comparable. Based on the above selection rules, the treatment group and the control group are defined in the following Table 30. Their geographical distributions are presented in Figure 14.

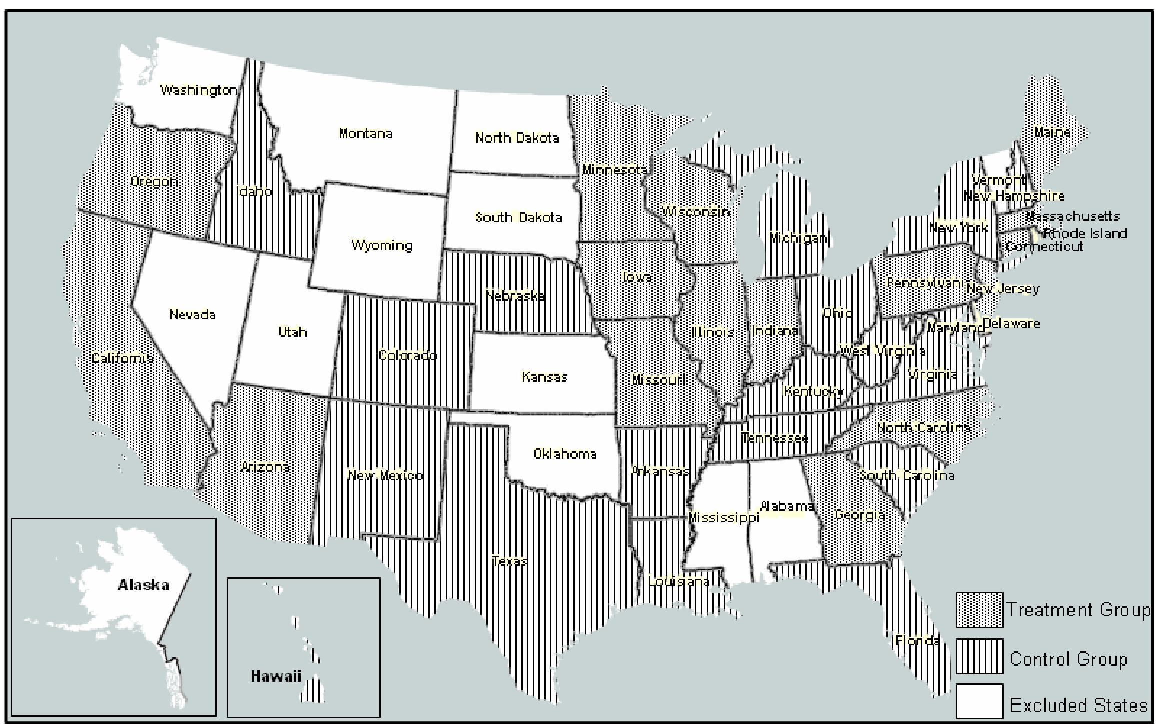


Figure 14. Geographical distribution of treatment and control states

Table 30. The definition of the treatment group and the control group

Criteria		Top ten states	Rest of states	Application
Single or double factor formula corporate tax	- R&D expense-based - General - Before 2000	California(1987-), Illinois(1990-2003), Massachusetts(1991-), New Jersey(1994-), Pennsylvania(1997-2006)	Arizona(1994-2003), Connecticut(1994-), Georgia(1998-2004), Indiana(1990-2013), Iowa(1985-), Maine(1996-), Minnesota(1981-), Missouri(1995-2004), North Carolina(1996-2006), Oregon(1989-2011), Wisconsin(1986-)	treatment
	- R&D expense-based - General - After 2000	Ohio(2001-) Texas(2000-2009)	Hawaii(2000-2010), Idaho(2001-2005), Louisiana(2003-2006), Maryland(2000-2006), New Mexico(2000-), South Carolina(2001-), West Virginia (2003-)	Control
	- R&D expense-based - Targeted - Before 2000		Arkansas(1997-), Colorado(1989-)	
	- R&D expense-based - Targeted - After 2000	Michigan(2002-)		
	- No R&D tax credit	New York	Florida, Kentucky, Nebraska, New Hampshire, Tennessee, Virginia	
Three-factor formula corporate tax	- R&D expense-based		Delaware(1999-2006), Kansas(1987-2000), North Dakota(1988-), Rhode Island(1994-), Utah(1999-2010)	Excluded
	- Employment-based		Mississippi(1989-), Oklahoma(1993-2003), Vermont(1992-1996)	
	- No R&D tax credit		Alabama, Alaska, District of Columbia	
Specific state corporate tax	- R&D expense-based - General or targeted	Washington(1995-2004)	Montana(1998-2010)	
No state corporate tax			Nevada, South Dakota, Wyoming	

For constructing the treatment group and the control group as described above, one important assumption, which is that spillover effects are prevalent within a region, is required. Figures 15 and 16 represent assumed scenarios of the DD setting and the DDD setting respectively. The same assumption is also applied to matching estimates because we utilize the matching method incorporating the DD method. These scenarios show us that all kinds of firms might be affected from state R&D tax credits either directly (represented as Firm A) or indirectly (represented as Firm B), and therefore all of these firms should be included in the firm-level analysis of this study. Based on the same assumption, the state-level aggregated data is also an alternative for testing the program effect of this study.

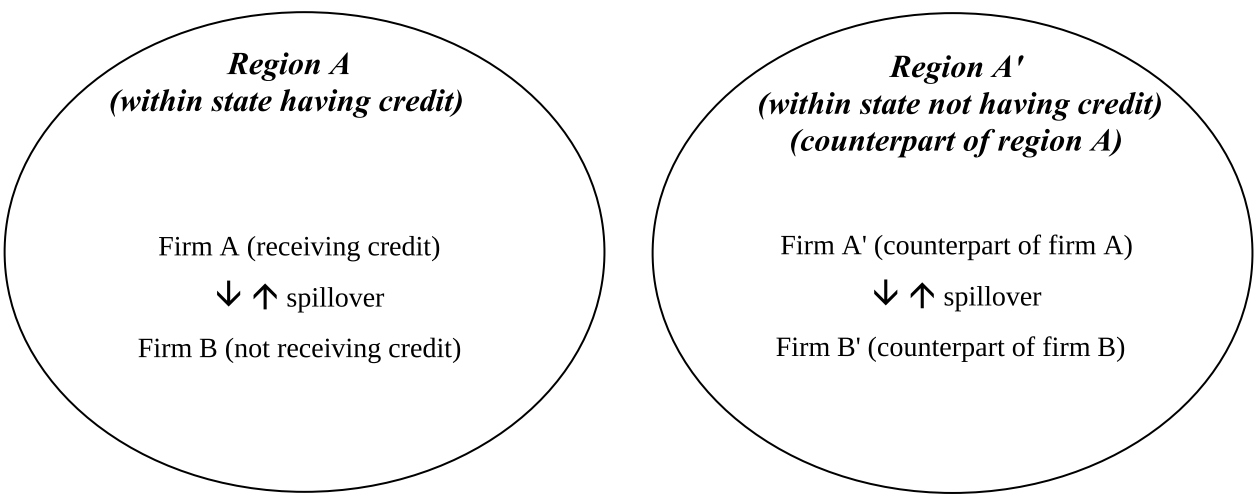


Figure 15. The scenario for the DD setting with spillovers

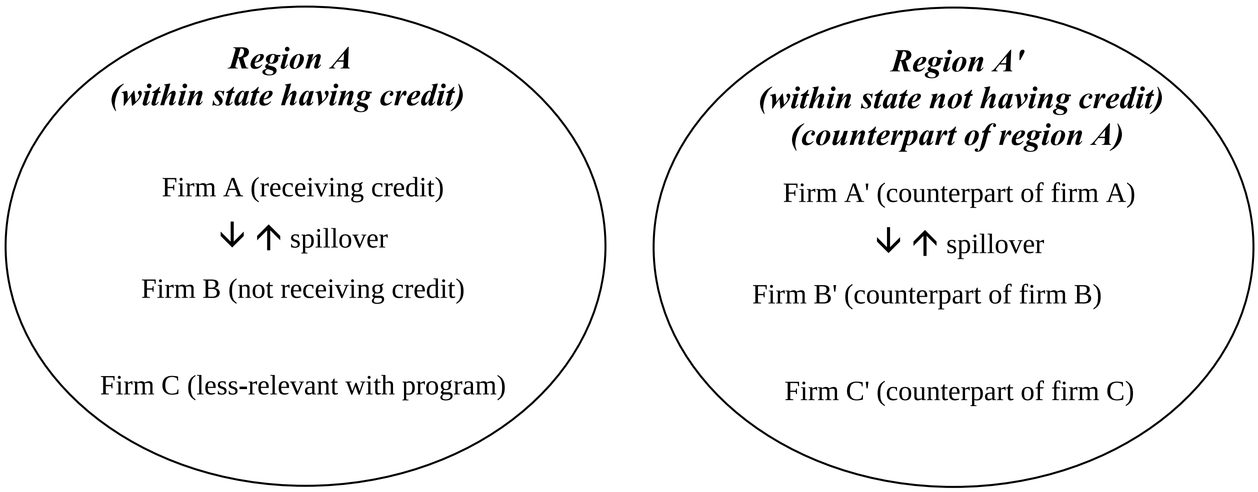


Figure 16. The scenario for the DDD setting with spillovers

In case of the DDD setting in Figure 16, the firms that are less-relevant with state R&D tax credit program (represented as Firm C) are constructed another control group for separating any possible effects from other factors. Because the less-influential group is defined as the 3-digit SIC level, this group can be recognized in both the firm-level and the state-level analyses depending on data availability.

Within a region having state R&D tax credits, there are firms receiving the credits and firms not receiving. In this case, if firms not receiving the credits don't derive any benefits from the credits, it is necessary to exclude these firms from the treatment group in this study. However, there are at least three bases for supporting these spillover effects, as discussed below.

First, many literatures supported spillover effects across R&D intensive firms at the regional level (see, for example, Romer, 1986, 1990; Lucas, 1988; Griliches, 1979, 1992; Aghion and Howitt, 1992 for the theoretical explanations and Jaffe et al., 1993; Almeida and Kogut, 1990, 1999; Audretsch and Feldman, 1996 for the empirical evidences).

Second, it is believed for obtaining spillover effects from firms receiving the credits, even firms not receiving the credits might make their own efforts such as increasing their R&D spending or employment in order to catch up the advanced R&D (see, for example, Cohen and Levinthal, 1989; Geroski, 1995; Klette, Møen and Griliches, 2000 for the theoretical explanations and Cohen and Levinthal, 1989; Branstetter and Sakakibara, 1998; Geroski et al., 1993 for the empirical findings).

Third, the tax credit law itself ensures the spillover effect by including contract research expenses. Based on this assumption, the effects are estimated as combined effects of "the direct effects to firms receiving the credits" and "the indirect effects to firms not receiving the credits but benefiting from spillovers." In particular, in the DDD setting, it is also assumed firms which are less-relevant to R&D tax credit are also less relevant to spillover effects and therefore they become a valid control group. The region, where geographically bounded spillover occurs, is defined either as a state for the state-level analysis or as a state and a county for the firm-level analysis.[94]

These scenarios can be modified in Tables 31 and 32 depicted below. Following the two possible scenarios with the existence of spillovers, the DD estimates consist of two possible effects, namely 1) the direct effect which means a difference of outcome changes of firms receiving credits and their counterparts and 2) the indirect effect which means a difference of outcome changes of firms not receiving credits but having a benefit through spillovers and their counterparts. The counterpart firms are defined as firms within states not providing credits but having similar characteristics to firms within states providing credits, and therefore these firms

[94] The state-level data and the county-level data are used as covariates or matching variables for the firm-level analysis depending on availability.

are supposed to either receive credits or have a benefit from spillovers if their states provide credits.

The DDD estimates consist of two possible effects which are the same as the DD estimates and one more possible effect which means outcome change of firms with a low level of R&D activities, and therefore are less-relevant with R&D credits. The third effect within the DDD estimates could capture outcome changes from other possible factors except the program, and therefore should be eliminated.

Table 31. The difference-in-differences formation

Outcome change* of Firm A – Outcome change* of Firm A'	=	Direct effect	(1)
Outcome change* of Firm B – Outcome change* of Firm B'	=	Indirect effect through spillover	(2)
Direct effect (1) + indirect effect (2)	=	Effect of program	

Notes: * Outcome change is defined as a difference of outcomes before and after program

Table 32. The difference-in-difference-in-differences formation

Outcome change* of Firm A – Outcome change* of Firm A'	=	Direct effect	(1)
Outcome change* of Firm B – Outcome change* of Firm B'	=	Indirect effect through spillover	(2)
Outcome change* of Firm C – Outcome change* of Firm C'	=	Outcome change from other possible factors	(3)
Direct effect (1) + indirect effect (2) – Outcome change from other possible factors (3)	=	Effect of program	

Notes: * Outcome change is defined as a difference of outcomes before and after program

6.3.3 Developing evaluation methods

Third, suggested evaluation methods, which are the DD/DDD method and the matching method incorporating the DD method, are examined and developed. These evaluation methods provide the magnitude of different outcome changes between the treatment group and the control group. In particular, the estimates become more precise by using the DDD method, which allows narrowing the treatment group to a subgroup strongly related to R&D tax credit. For this estimation, the influential group A consists of the high-technology industry and the less-influential group B is composed of the less R&D-intensive industries with the assumption that state R&D tax credits make strong effects on R&D-intensive industries.

Another way to deal with rival hypotheses by using the DD/DDD estimation is adding covariates which are relevant to outcome changes. The selected covariates are 1) the lagged GSP (for state) or the lagged sales (for firm) for capturing primary factors to decide future R&D investments as well as the level of employment, 2) civilian labor force for capturing the overall capacity of regional human capital, 3) knowledge stocks for measuring the level of R&D, 4) unemployment rate for capturing general economic conditions especially for the analysis of R&D spending,[95] 5) private R&D spending especially for the analysis of employment, and 6) federal funding as the direct government intervention especially for the analysis at the state level.[96] The relationship between these selected covariates and above five rival hypotheses is shown in Table 33. It indicates the possible observational levels and data sources for the selected covariates as well.

---

[95] Unemployment rate is not included in the analysis of employment due to possible endogeneity.

[96] Federal funding is constructed at the state level. Therefore it is not comparable with the private R&D funding which is constructed at the firm level and therefore this variable is excluded in the firm level analysis.

Table 33. The selected variables to deal with rival hypotheses as covariates and matching variables

Rival hypothesis		Evaluation strategy	Factors/variables
Five environments	Detailed features		
Different state tax structures	- Corporate tax system - State R&D tax credit	- Constructing valid control group	- State law
Different business environments	- Human capital - Economic condition - Investment capacity - Agglomeration economies	- Controlling from observable variables DD/DDD method with covariates State level	- Gross State Product/Personal Income - Civilian labor force - Unemployment rate
		- Controlling from observable variables Matching method State level/county level	- Establishment density - Per capita income - Poverty rate - Population change - Crime rate
		- Controlling from unobservable variables	- Time-specific, cross-sectional fixed-effects
Different R&D Environments	- Possibility of spillovers - Competition - Level of R&D - Knowledge stock - R&D performance of universities and research institutions	- Controlling from observable variables DD/DDD method with covariates State level/county level	- Knowledge stock from patent data
		- Controlling from observable variables Matching method State level	- Establishment density - University funding - Share of labor force over college education and high school education
		- Controlling from unobservable variables	- Time-specific, cross-sectional fixed-effects
Different government Interventions	- Direct funding - Indirect Tax incentives - Other economic development strategies	- Controlling from observable variables DD/DDD method with covariates State level	- Federal funding
		- Controlling from observable variables Matching method State level	- SBIR grants
		- Controlling from unobservable variables	- Time-specific, cross-sectional fixed-effects
Different firm characteristics	- Investment capacity - Firm size - Establishment status	- Controlling from observable variables DD/DDD method with covariates Firm level	- Sales - Number of employees
		- Controlling from observable variables Matching method Firm level	- Firm size
		- Controlling from unobservable variables	- Time-specific, cross-sectional fixed-effects

Within the DD method or the DD method, the observations for covariates should be available by year as same as the observations for outcome variables for reducing the systematic differences of the comparison groups. Then, by applying the fixed-effect panel analysis, possible effects from unobservable variables could be eliminated as well. In sum, the different business environment, the different R&D environment, the different government policy environment and the different firm characteristics, as the rival hypotheses under the section 6.2.2~6.2.5, are ruled out by controlling possible outcome changes of the control group, possible outcome changes from covariates, and possible outcome changes from time-specific effects and cross-sectional fixed-effects. In addition, this paper estimates White-corrected standard errors for obtaining the heteroskadasticity-consistent standard errors (White, 1980).

Along with the DD method or the DDD method, the matching method provides alternative estimates for policy impact. The matching method incorporating the DD method is adopted in this study based on the suggestion of that the DD estimates are allowed to absorb possible effects from unobservable variables. Therefore, the dependent variables for the matching method are defined as the average changes of R&D spending and employment. In addition, multiple matching metrics are used for finding multiple evidences and they are the Mahalanobis metrics and the propensity score metrics. Within the propensity score metrics, we also use both the nearest neighbor matching based on one-to-one matching and multiple matching by specifying the number of matches. Next, for defeating several rival hypotheses mentioned above, variables for the proper matching are suggested from different levels of regions. Here the data requirement for matching variables is less strict than covariates, in the sense that the observations for the matching should capture the comparable degree of observable characteristics but, they are not necessarily at the every year base. Therefore, the possible matching variables could encompass every covariate and also be extended. Indeed, matching variables can be chosen with less-strict requirements because this method is one of non-parametric method. This possible diversity could provide the better condition for the better matching. The base year for comparable degree of observable characteristics is chosen by the year of 1990 because most of state R&D tax credits were enacted after 1990.

The sets of matching variables are constructed in two ways, which are 1) selected variables, which are same as covariates and 2) all selected matching variables as follows. Possible effects from different business environments (from rival hypothesis 6.2.2.) are considered from the following variables: 1) GSP, 2) labor force, 3) per capita income, 4) unemployment rate, 5) poverty rate 6) population change and 7) crime rate. Next, possible effects from different R&D environments (from rival hypothesis 6.2.3) are eliminated from the selected variables, which are 1) the level of R&D as knowledge stocks, 2) university funding, 3) private R&D expenditure, 4) establishment density, and 5) share of labor force over college education and high school. Then, different policy environments (from rival hypothesis 6.2.4.) are controlled by 1) SBIR grants and 2) federal funding. Finally different firm characteristics (from rival hypothesis 6.2.5) are captured by the following variables that are 1) sales, 2) R&D expenditure, and 3) firm size.

The extensive set of covariates and matching variables with the corresponding rival hypotheses is summarized at the above Table 33. Based on the level of analysis and data availability, each matching variable is constructed at different levels of regions as well as different analytical levels.

6.3.4 Empirical specifications

Based on the above discussion for better evaluations, the detailed estimates are specified below. These specifications include four types of estimates under the DD/DDD methods, which are 1) the simple DD estimates, 2) the DD estimates with covariates, 3) the simple DDD estimates, and 4) the DDD estimates with covariates, and two types of estimates under the matching methods incorporating the DD method, which are 1) the matching estimates based on the Mahalanobis metrics and 2) the matching estimates based on the propensity score metrics.

6.3.4.1 The DD/DDD methods

First the simple DD estimates are specified. Within this specification, we have only dummy variables as the combined form of pre/post program dummy and treatment/control dummy (T·Post=1, if post-program and treatment), which indicates whether there is a difference of outcome change between the treatment group and the control group. If the estimate (i.e. $\delta(D_1D_2)$) is positive, it means that the outcome change of the treatment group is larger than that of the control group and the amount of difference corresponds to the effect of the program. By doing fixed-effect panel analysis, the state (or industry specific effect and the year specific effect are absorbed. Here, outcome variables, as dependent variables, is defined as yearly R&D spending and employment of state (or firm).

$$Y_{ist} = A_s + B_t + C_i + \delta D_{st} + \varepsilon_{ist} \text{ -- } (3)$$

Y_{ist} : presumed outcomes from the program of state (or firm): R&D spending and employment
A_s : fixed effects for state
B_t : fixed effects for year
C_i : fixed effects for industry
D_{st} : dummies for measuring program effects – post-program dummy (D_1), treatment dummy (D_2), post-program · treatment dummy (D_1D_2)
ε_{ist} : disturbance term
δ_1 (D_1D_2) : coefficient for program effect

Second, the DD estimates with covariates are specified. This specification can make DD estimates more precise by capturing other possible effects from observable variables, which are identified in rival hypotheses 6.2.1 through 6.2.5, such as general economic conditions (from unemployment rate), R&D investment capacity (from lagged GSP/personal income at the state level and lagged sales at the firm level) and overall regional human capital (from civilian labor force), the regional level of R&D (from knowledge stock) and government direct intervention (from federal funding), the previous performance (from the lagged R&D expenditure and the lagged employment). This specification can also capture other possible effects from unobservable variables by using the fixed-effect panel analysis.

$$Y_{ist} = A_s + B_t + C_i + \beta X_{ist} + \delta D_{st} + \varepsilon_{ist} \text{ -- } (4)$$

Y_{ist} : presumed outcomes from the program of state (or firm): R&D spending and employment
A_s : fixed effects for state
B_t : fixed effects for year
C_i : fixed effects for industry
X_{ist} : covariates – lagged GSP/lagged sales, civilian labor force, knowledge stock, lagged private R&D expenditure, lagged employment (only for the analysis of employment) federal funding (only for the analysis at the state level), and unemployment rate (only for the analysis of R&D spending)
D_{st} : dummies for measuring program effects – post-program dummy (D_1), treatment dummy (D_2), post-program · treatment dummy (D_1D_2)
ε_{ist} : disturbance term
$\boldsymbol{\beta}$: coefficients for relationships between covariates and outcome
δ_1 (D_1D_2) : coefficient for program effect

Third, the simple DDD estimates are specified. Within this specification, treatment group narrows down R&D-intensive industry within states having R&D tax credits. For doing this, the dummy variable for R&D-intensive/non-R&D-intensive industry is added and therefore, coefficient of dummy variable for R&D intensive, post-program and treatment (D_1=1, D_2=1 and D_3=1) corresponds to the program effect. Following the above the DD estimates the fixed-effect panel analysis is performed. The estimates of R&D spending at the state level are excluded due to data availability.

$$Y_{ist} = A_s + B_t + C_i + \delta D_{st} + \varepsilon_{ist} \quad \text{--} \quad (5)$$

Y_{ist} : presumed outcomes from the program of state (or firm): R&D spending and employment
A_s : fixed effects for state
B_t : fixed effects for year
C_i : fixed effects for industry
D_{st} : dummies for measuring program effects – post-program dummy (D_1), treatment dummy (D_2), R&D intensive industry dummy(D_3), R&D-intensive industry · post-program dummy (D_1D_2), R&D-intensive industry · treatment dummy (D_1D_3), post-program · treatment dummy (D_2D_3), R&D-intensive industry · post-program · treatment dummy (D_1D_2 D_3)
ε_{ist} : disturbance term
δ_1 (D_1D_2 D_3) : coefficient for program effect

Fourth, the DDD estimates with covariates are specified. In this specification, the DDD estimates might be more precise because other relevant effects except the program can be captured by adding some covariates as above specification (2). Especially, based on the assumption of that overall economic conditions and previous economic performances also affect the decisions of the level of R&D investments and employments, several covariates, such as unemployment rate, lagged GSP/personal income/sales, civilian labor force, knowledge stock, federal funding, and the lagged R&D expenditure/employment are chosen. Again, this specification is not available for the estimates of R&D spending at the state-level analysis because the industry-level data is not available.

$$Y_{ist} = A_s + B_t + C_i + \beta X_{ist} + \delta D_{st} + \varepsilon_{ist} \quad \text{--} \quad (6)$$

Y_{ist} : presumed outcomes from the program of state (or firm): R&D spending and employment
A_s : fixed effects for state
B_t : fixed effects for year
C_i : fixed effects for industry
X_{ist} : covariates – lagged GSP/lagged sales, civilian labor force, knowledge stock, lagged private R&D expenditure, lagged employment (only for the analysis of employment) federal funding (only for the analysis at the state level), and unemployment rate (only for the analysis of R&D spending)
D_{st} : dummies for measuring program effects – post-program dummy (D_1), treatment dummy (D_2), R&D-intensive industry dummy (D_3), R&D-intensive industry · post-program dummy (D_1D_2), R&D-intensive industry · treatment dummy (D_1D_3), post-program · treatment dummy (D_2D_3), R&D-intensive industry · post-program · treatment dummy (D_1D_2 D_3)
ε_{ist} : disturbance term
$\boldsymbol{\beta}$: coefficients for relationships between covariates and outcome
δ_1 (D_1D_2 D_3) : coefficient for program effect

6.3.4.2 The matching method

The analysis by using the matching method is conducted only at the firm level because a substantially large number of datasets are required for conducting this method. First, the matching method incorporating the DD method based on the Mahalanobis metrics is specified. In this specification, paired matches are constructed by comparing the unique Mahalanobis distance of matching variables which are listed in Table 33. In particular, the nearest neighbor matching estimator based on the one-to-one matching is estimated. The estimates are provided by comparing the paired differences of outcome changes and we also allow the multiple matching with replacement for better estimation. As mentioned above, the matching variables are much numerous than covariates based on less strict data requirements. The sets of matching variables include 1) GSP (or personal income), 2) labor force, 3) per capita income, 4) unemployment rate, 5) poverty rate, 6) population change and 7) crime rate, from rival hypothesis 6.2.2. Next, from rival hypothesis 6.2.3, 1) the level of R&D-knowledge stock, 2) university funding, 3) private R&D expenditure, 4) establishment density, and 5) share of labor force over college education or high school are selected. Then, from rival hypothesis 6.2.4, 1) federal funding and 2) SBIR grants are chosen. Next, 1) sales and 2) firm size are chosen from rival hypothesis 6.2.5. The base year for matching variables is chosen as the year of 1990.

Following the DD estimation method, the dependent variables are defined as the average changes of economic outcomes (i.e. R&D spending and employment). The difference is that the dependent variable itself is constructed as the average change before and after program from the multiple time-series datasets. Indeed, the bootstrapping method is used for assessing the significance of estimators. All the analyses utilizing the matching method are performed by using the *psmatch2* STATA procedure developed by Leuven and Sianesi (2003).

Second, the matching method incorporating the DD method based on the propensity score metrics is specified. In this specification, we use the propensity score for comparing the closeness of observations instead of the Mahalanobis distance. Within this method, I estimate two matching estimators, which are the nearest neighbor matching estimator based on one-to-one matching with caliper, and the nearest neighbor matching estimator based on multiple matching (# of matches: 3) with caliper. Other specifications are same as the above specification for the matching method incorporating the DD method based on the Mahalanobis metrics.

In this section, I examined the detailed empirical specifications by two selected statistical methods, which are the DD/DDD methods and the matching method incorporating the DD method. In the next section, I examine the data in detail in terms of where they are obtained, what important characteristics and features they have, and how they are modified and constructed for the empirical analysis.

6.4 DATA

In this section, I examine the data to be used. The data to be used in this study is broadly categorized by three, as dependent variables, covariates, and matching variables.

The dependent variables, which are historical observations for R&D spending and employment, comes from two different datasets for the state level analysis but from the same datasets for the firm level analysis. Based on the different data sources, these variables are originally constructed in a different way. The observations for R&D spending at the state level are only biannual ones and do not break down by industry, while other dependent variables are annual observations by the 3-digit SIC level.

The covariates, for capturing possible systematic differences of the treatment states and the control states within the DD/DDD specifications, are obtained at the various levels, namely, the state level, the industry level and the firm level depending upon data availability. Because the yearly observations are necessary for these covariates, the number of selected variables is limited at each level of analysis. However, in case of the matching variables, the data requirement is less strict in the sense that these variables should capture the characteristics of preprogram but not necessarily based on yearly observations. Therefore, all covariates are included in the matching variables and some variables are added. The primary sources of these variables are summarized in Table 34 and the important characteristics and features of each variable follow.

Table 34. The data sources of variables to be used in this study

Variable name	Application: Category of variables	Application: Level of analysis	Observational level	Data source
R&D spending	Dependent variables	State	State	Total funds for industrial R&D performance by the NSF
	〃	Firm	Firm	Compustat by Standard & Poor’s
Number of Employees by industry/state	〃	State	State/industry	Current Employment Statistics (CES) by the Bureau of Labor Statistics (BLS)
		Firm	Firm	Compustat by Standard & Poor’s
Gross State Product (GSP)	Covariates & matching variables	State	State/industry	The GSP by the Bureau of Economic Analysis (BEA)
Personal Income (PI)	〃	State	State/industry	The PI by the Bureau of Economic Analysis (BEA)
Net sales	〃	Firm	Firm	Compustat by Standard & Poor’s
Unemployment rate	〃	State/firm	State	Local Area Unemployment Statistics (LAUS) program by the BLS
Civilian labor force	〃	State	State	Current Population Survey (CPS) by the Bureau of Census
Number of patent	〃	State/firm	State/county	The patent data by the USPTO and Hall, Jaffe and Trajtenberg (2001)
Federal funding	〃	State	State	The IRIS by the SRS of the NSF
Per capita income	Matching variables	Firm	County	The City and County Data Book (88, 94) by the Bureau of Census
Poverty rate	〃	Firm	County	The City and County Data Book (88, 94) by the Bureau of Census
Population change	〃	Firm	County	The City and County Data Book (88, 94) by the Bureau of Census
Crime rate	〃	Firm	County	The City and County Data Book (88, 94) by the Bureau of Census
Establishment density	〃	Firm	State	The Economic Census (82, 87, 92) by the Bureau of Census
Share of labor force over college or high school education	〃	Firm	State	The Decennial Census (80, 90) by the Bureau of Census
University funding	〃	Firm	State	The webCASPAR database by the NSF
SBIR grants	〃	Firm	State	The SBIR data from the SBA

Private R&D spending by state, which is the primary dependent variable of the state level analysis, is obtained from the database of “the Industrial Research and Development Information System (IRIS),” provided by the Division of Science Resources Statistics (SRS) of the National

Science Foundation (NSF).[97] This database is constructed based on the Survey of Industrial Research and Development by the Bureau of the Census since 1957. The data are collected from publicly or privately held R&D-performing companies by a sample survey. The above database contains biannual observations between 1963 and 1998, which is "total funds for industrial R&D performance, by state for selected years: 1963–98," defining private R&D spending as the portion of the R&D expenditure performed by industrial firms and financed by their own funds. In addition, the observations for 1999 are added from the annual report by the SRS, which is "Research and Development in Industry: 1999" (NSF, 2002). Accordingly, there are the 20-year observations.

Employment at the state level, as the other economic outcome variables and dependent variable in this study, is acquired from the Current Employment Statistics (CES) by the Bureau of Labor Statistics (BLS).[98] This dataset is collected from an extensive sample including about 160,000 businesses and government agencies from about 400,000 individual worksites, which constitutes to approximately one-third of the whole non-farm payroll workers. Accordingly, this dataset contains detailed observations classified by the 3-digit SIC level, based on the monthly collections from 1953 for most states.[99] These detailed observations allow identifying the high-technology industry and providing rich observations. This paper uses the seasonally adjusted annual average data from 1963,[100] which include the 38-year observations.

---

[97] This database is accessible via http://www.nsf.gov/statistics/iris/.

[98] This database is accessible via http://www.bls.gov/sae/home.htm.

[99] In recent data, 4-digit SIC level data is included and also metropolitan area level (MSA level) data is also included for the details.

[100] The year of 1963 is chosen for the initial year for observations because R&D spending data is available from this year.

For constructing covariates, the total Gross State Product (GSP) provided by the Bureau of Economic Analysis (BEA) is used for measuring the overall economic capacity by region.[101] By definition, the GSP is "the value added in production by the labor and property located in a state" and measured by "industry's gross output, including sales or receipts and other operating income, commodity taxes, and inventory change, minus its intermediate inputs, including consumption of goods and services purchased from other U.S. industries or imported" (Aman, Downey, and Panek, 2005, p.82). The GSP is available from 1977. Alternatively, the Personal Income (PI) provided by the BEA, can be used for controlling the difference of overall economic capacity.[102] By definition, the PI is "the income received by all persons from all sources" and measured by "the sum of net earnings by place of residence, rental income of persons, personal dividend income, personal interest income, and personal current transfer receipts" (the BEA, 2006, p.5). The PI is available from 1969. As other covariates, the annual unemployment rate by state comes from the Local Area Unemployment Statistics (LAUS) program by the BLS,[103] and the yearly civilian labor force is provided by the Current Population Survey (CPS) conducted by the Bureau of Census for the BLS.[104] Both are available from 1971. The unemployment rate variable is excluded in the empirical specification for measuring effects on employment for avoiding the possible endogeneity.

In addition, the different policy environment at the federal level could be captured by the federal-level R&D expenditures provided by the IRIS of the SRS in the NSF, which is the same data source with private R&D spending. Besides the federal-level R&D expenditures, the R&D

---

[101] This database is accessible via http://www.bea.gov/bea/regional/data.htm.

[102] This database is accessible via http://www.bea.gov/bea/regional/data.htm.

[103] This database is accessible via http://www.bls.gov/lau/.

[104] This database is accessible via http://www.bls.census.gov/cps/cpsmain.htm.

expenditures performed by industrial firms but financed by local governments or university/college are another possible sources of funds, but these are not available in this database, especially at the state level. The federal-level R&D expenditures are defined as the portion of R&D expenditures performed by industrial firms and totally financed by federal government (NSF, 1999). Following this definition, the federal government R&D expenditures for its own R&D performance are excluded. So, there is a difference of the total federal government R&D funds between in this dataset (the IRIS by the NSF) and in "Federal Funds for Research and Development" by the NSF. In this study, private R&D activities, in detail how the private companies react for deciding their R&D expenditures depending on the existence of the R&D tax credit, are primary concern, and therefore we particularly interested in this specific private R&D activities, not a whole R&D activities. Consequently, the portion of federal government R&D expenditures specially supporting for private R&D activities is the more appropriate measurement which is related to private R&D expenditures for its own R&D activities and therefore the data of the IRIS is chosen for this study. Meanwhile, it should be noted that there is a possibility of connection of the portion of R&D activities by government itself with private R&D through knowledge spillovers, but this relationship differs from the one mentioned above.

The knowledge stock by state for measuring the different level of R&D is constructed as the five-year sum of number of patents assigned.[105] The patent data is provided by the United States Patent and Trademarks Office (USPTO) and Hall, Jaffe and Trajtenberg (2001). Under the assumption of that the knowledge could be accumulated with the certain obsoleteness, the five-year numbers of patents are aggregated with the depreciation rate as 10%, which is

[105] There are some other possible measurements for level of R&D and knowledge stock.

suggested in previous studies (See, Smith, 1999 and Caballero and Jaffe, 1993). The above five variables, which are the GSP/PI (comparable), the unemployment rate, federal funding, and knowledge stocks, are met the necessity of annual observations for covariates in empirical specifications.

Next, the firm-level data is obtained from the Compustat by the Standard & Poor's. This dataset contains yearly observations of various variables such as R&D spending, the number of employees, and sales. The observed years are 42 years from 1960 to 2001 and the observed companies are approximately 23,000 in total and about 6,400 with their R&D spending. Each firm classifies 6-digit industry level by using the North American Industry Classification System (NAICS) and its location is identified by state and county. The NAICS code is converted by the 3-digit SIC code for identifying the high-technology industry.

Then, the sources of matching variables are examined. The observation year for matching variables, which is necessary for comparing the degree of observable characteristics prior to the program, is selected as 1990 which is an average year that the law was adopted across states. The matching variables come from two primary data sources, which are the County and City Data Books and the Economic Census, provided by the Bureau of Census.[106] From the County and City Data Books which are the collections of several economic variables from various sources in selected years, it is able to capture the different economic conditions across regions for ruling out rival hypotheses. These variables are the population change, the share of labor force over college education, the share of labor force over high school education, and the crime rate. I use the County and City Data Books published in 1988 and 1994, and the

[106] These databases are accessible via http://fisher.lib.virginia.edu/collections/stats/ccdb/ and http://www.census.gov/econ/census02/ respectively.

original data sources for these data collections are the Decennial Census of Population and Housing in 1990, and the Uniform Crime Reporting (UCR) Program by U.S. Federal Bureau of Investigation in 1990 and 1991. From the Economic Census performed in the quinquennial years, the information for the number of establishments is obtained. I measure the establishment density as the number of establishments divided by population. Along with the above variables, the Per Capita Income in 1989, provided from the Bureau of Census and based on the Decimal Census,[107] are also used in this study.

In addition, the yearly R&D expenditure of universities and colleges is obtained from the database of "the Integrated Science and Engineering Resource Data System (webCASPAR)", provided by the Division of Science Resources Statistics (SRS) of the National Science Foundation (NSF).[108] This data is used for capturing possible spillover effects from R&D activities of universities and colleges. This data is available from 1972 to 2001 and constructed by state. The Small Business Innovation Research (SBIR) funds, for capturing different R&D environment and policy environment, are provided by the Technology Resources Network (Tech-*net*) of the Small Business Administration (SBA). This database provides the detailed information on SBIR awards, categorizing by the name of company, firm size, firm location by city and state, ownerships, the award phases and so on.[109] The SBIR funds specially target the entrepreneurial companies, which are small (limited 500 employees) and highly innovative, for supporting their researches and their commercialization. This program started at the year of

---

[107] This database is accessible via http://www.census.gov/hhes/www/income/histinc/histinctb.html.

[108] This database is accessible via http://webcaspar.nsf.gov/index.jsp;jsessionid=7784118BC68AC63AFA67B6145B29F7B4?subHeader=WebCASPARHome&showHelp=false.

[109] It also provides the similar information for Small Business Technology Transfer (STTR) awards, Advanced Technology Program (ATP) awards, and Manufacturing Extension Partners (MEP) centers. This database is accessible via http://tech-net.sba.gov/.

1982 under the Small Business Innovation Development Act. I construct the SBIR funds as the total amount of funds by state between 1983 and 1990.

For finding the similar economic condition at the firm level, I use the county-level data if available, or instead the state-level data. In case of covariates, the county-level data at the every year base is mostly not available and therefore the state-level data is replaced. Under this construction, the federal funding variable is excluded from covariates in the empirical specification for measuring effects on R&D spending and employment at the firm level because private funding at the firm level and federal funding at the state level is not comparable. In case of matching variables, some variables, such as civilian labor force, unemployment rate, the number of establishment, the share of labor force over college education, the share of labor force over high school education, the per capital income, the population change, and the crime rate, are available at the county level. However, other variables, such as SBIR funds, the R&D expenditure of universities and colleges, and federal funding, are constructed at the state level.[110]

In this study, outcome variables which are R&D spending and employment as dependent variables and other covariates for DD/DDD method are measured by the level, the estimates in which indicate the amount of outcome changes, and transformed in logarithm, the estimates in which indicate the % outcome changes. Indeed, the 1982-84 base Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W), published by the BEA, has been used to deflate all dollar value variables.[111]

---

[110] In case of the SBIR funds and university funds, the city-level data is available. However, the definition of the city and that of county is not comparable, in other words, their boundaries cross each other. So, in this paper, the city-level data is excluded.

[111] Most dollar value variables are collected in current dollar however, there are some exceptions. University funding data is originally constructed in 1996 constant dollar by using GDP Implicit Price Deflators, and therefore this data is converted into current data and then reconverted into the 1982-84 constant dollars.

Tables 35 and 36 represent the descriptive statistics of the data to be used in this research, at the state level and at the firm level respectively. Those statistics are constructed by treatment and control groups.

Table 35. The descriptive statistics of the data for the state-level analysis

Data			# of observations	Mean	Standard error	Minimum	Maximum
Dependent variable	Private R&D spending (millions)	Treatment states	220	2,085.6	2,912.3	12.9	21,449.7
		Control states	199	1,464.8	1,917.6	5.5	10,772.0
	Employment (thousands)	Treatment states	486	427.9	474.1	3.8	2,412.9
		Control states	554	309.1	470.6	2.0	1,679.6
Covariates	Federal R&D funding (millions)	Treatment states	203	829.5	1,713.1	0.9	10,117.8
		Control states	177	509.7	457.0	0.6	2,395.8
	GSP (millions)	Treatment states	221	123,725.2	123,961.5	12,249.4	743,477.3
		Control states	247	98,150.9	101,448.2	10,431.8	455,804.5
	Personal income (millions)	Treatment states	340	90,020.1	92,114.9	6,606.1	609,596.9
		Control states	380	70,003.5	76,689.4	4,808.2	377,391.7
	Laborforce (thousands)	Treatment states	262	3,384,308	2,886,117	478,000	16,600,000
		Control states	282	2,799,712	2,469,484	399,000	10,200,000
	Unemployment rate (%)	Treatment states	262	6.02	1.95	2.5	11.8
		Control states	282	6.26	2.24	2.5	18.0
	Knowledge stock (tens)	Treatment states	306	4,779.2	5,073.1	179.9	26,601.1
		Control states	342	2,864.1	3,651.0	61.0	16,408.8

Table 36. The descriptive statistics of the data for the firm-level analysis

Data			# of observations	Mean	Standard error	Minimum	Maximum
Dependent variable	Private R&D spending (millions)	Treatment	29842	19.7	83.4	0	3189.7
		Control	16168	29.1	212.3	0	5775.4
	Employment (thousands)	Treatment	41625	4.7	14.5	0	238
		Control	29497	6.2	33.9	0	876.8
Covariates -firm level	Netsales (millions)	Treatment	43450	428.6	1371.7	0	35534.9
		Control	30353	638.3	3775.3	0	110394
Covariates -county level	Knowledge stocks (tens)	Treatment	3546	348.1	1156.2	0.33	12988.2
		Control	2859	314.7	919.4	0.34	9164.1
Covariates -2digit SIC level	Laborforce (thousands)	Treatment	150	173,006	179,013	13,959	786,555
		Control	149	170,694	132,552	42,363	530,323
Matching variables -county level	Per capita income	Treatment	1099	9572.8	2163.9	4403.7	23149.2
		Control	1374	8939.5	2249.2	3386.6	22725.9
	share of high school graduates	Treatment	1100	71.4	9.1	42.8	91.9
		Control	1416	67.4	10.9	31.6	95.5
	Share of college graduates	Treatment	1100	13.5	6.5	3.7	46.1
		Control	1416	13.2	6.9	3.7	53.4
	# of crime	Treatment	1100	2960.9	2259.2	0	16031
		Control	1416	3249.8	2359.6	0	20179
	Poverty rate	Treatment	1100	14.1	6.1	2.2	48.7
		Control	1416	17.8	8.1	0	60
	Manufacturing establishment density	Treatment	1097	0.001	0.0007	0.00006	0.006
		Control	1396	0.001	0.0007	0.0001	0.007
Matching variables -state level	SBIR funds	Treatment	17	12.9	23.7	0.5	89.7
		Control	19	6.4	6.6	0.1	19.5
	University funds	Treatment	17	3435.6	3356.0	168.8	14742.6
		Control	19	2493.7	2886.3	252.8	10552.5

6.5 SUMMARY

Before going further, analytical methods and estimators which are used in this study are summarized for clarifying. Because multiple methods and estimators are suggested for finding robust evidences, it is necessary to summarize these methods and estimators for organizing the next analysis chapter for better understanding the results.

Table 37 illustrates the summary of estimators to be analyzed in this study. I use two different regional level datasets, which are one at the state level and the other at the firm level. These datasets are mostly constructed by the 3-digit SIC level.[112] In each level, there are two dependent variables – R&D spending and employments – which are assumed to be the major outcomes of R&D tax credits. Based on methodological necessity and the particular interest in the high-technology industry, each outcome is assessed into two categories, which are the all industry (defined as manufacturing and service industries) and the high-technology industry. By using two different regional level data as well as multiple dependent variables, we aim to provide more reliable and robust evidence of program effects based on multiple methods and multiple estimates. With the overall assessment of each outcome, we also examine the effect of each outcome by firm size by using the firm-level data. By doing so, we can test a questionable belief in different effects by firm size and especially in the existence of positive program effect for small firms.

[112] The exception is R&D spending at the state level, which is total industry R&D spending of state.

Table 37. The summary of estimators

State level	R&D spending	All industry	Overall
	Employment	All industry	Overall
		High-technology industry	Overall
Firm level	R&D spending	All industry	Overall
			Small, medium, large
		High-technology industry	Overall
			Small, medium, large
	Employment	All industry	Overall
			Small, medium, large
		High-technology industry	Overall
			Small, medium, large

Then, the analysis is performed from two major analytical methods – the difference-in-differences method and the matching method – which are chosen from the above program evaluation strategy section. As mentioned above, the difference-in-differences methods could provide the elaborate estimates through difference-in-difference-in-differences methods and DD/DDD methods provide the precise estimator for program effect by adding covariates. Accordingly, among possible four kinds of the DD estimators, two estimators with covariates are discussed in the next analysis chapter as shown in Table 38. The DD method is applied to both analyses for the all industry and the high-technology industry, and the DDD method is applied to the analysis for the high-technology industry level.

In case of the matching method, the nearest matching method is applied by using two different kinds of metrics, one of which is the Mahalanobis metrics and the other of which is the propensity score metrics. In the nearest neighbor matching method by using Mahalanobis metrics, the simple matching estimate, which is one-to-one matching without caliper, is estimated. In the nearest neighbor matching method by using propensity score metrics, two estimates are estimated as follows: 1) one-to-one matching with caliper, 2) multiple matching (# of matches: 3) with caliper. Indeed, by applying the matching method incorporating with the DD method, we can find the more accurate estimators by controlling the unobserved variables as well. About matching variables, the set of matching variables which are suggested from five categories of rival hypotheses are used. The significance of the matching estimate is judged by using the bootstrapping method. The matching method is conducted only at the firm level analysis due to the necessity of substantially large number of observations. Table 38 illustrated below is the summary of analytical methods and their applications in this study.

Table 38. The summary of analytical methods

<table>
<tr><th colspan="3">Analytical methods</th><th>Application</th></tr>
<tr><td rowspan="2">Difference-in-differences (DD) method</td><td colspan="2">Difference-in-differences (DD) method with covariates</td><td>state level, firm level
all industry, high-technology industry</td></tr>
<tr><td colspan="2">Difference-in-difference-in-differences (DDD) method with covariates</td><td>state level, firm level
high-technology industry</td></tr>
<tr><td rowspan="3">Matching method with DD method
: nearest neighbor matching</td><td>Mahalanobis metrics</td><td>One-to-one matching without caliper</td><td>firm level
all industry, high-technology industry</td></tr>
<tr><td rowspan="2">Propensity score metrics</td><td>One-to-one matching with caliper</td><td>firm level
all industry, high-technology industry</td></tr>
<tr><td>Multiple matching (# of matches: 3) with caliper</td><td>firm level
all industry, high-technology industry</td></tr>
</table>

With the above multiple estimators, we judge the effectiveness of the credits based on the existence of significant estimates. Importantly, the effects which are estimated in this study as the difference of average outcome changes include the direct effects to firms which receive the credits, and the indirect effects to firms which do not receive the credits. The possible indirect effects include effects from sub-contract, spillover effects, or catching-up efforts.

Then, I examine evaluation strategies developed in this study by assessing how these strategies deal with rival hypotheses that are defined in the section 6.2. Rival hypotheses are defined under five environments in Table 27 and are eliminated by constructing valid control groups and reducing systematic differences caused from these factors through covariates and matching variables and time-specific and cross-sectional fixed effects. However, some factors in rival hypotheses are still remaining due to the lack of measurable variables. In particular, in the DD/DDD methods, more factors in rival hypotheses are possibly remaining than in the matching methods due to the more limited availability of data for covariates than matching variables.

Table 39 summarizes the evalution strategies, unexplained parts in the model and overall assessment for each rival hypothesis in this study. In this table, three highlighted factors, which are possible different level of agglomeration economies, the level of competition and available human capital, are recognized as the fundamental factors for deciding the level of private R&D activities especially at the regional level but not being able to be eliminated directly within the model. These factors are relevant to selection (i.e. systematic differences in characteristics between experimental observations and others) among the possible threats of validity. Because private R&D activities are quite skewed across states and across regions, the outcome changes from these possible different environments might combine the outcome change from program and cause to evoke the confusion.

Table 39. Re-evaluating rival hypotheses

Rival hypotheses		DD/DDD methods	Matching method	Possible remaining parts	Overall Assessment
Different state tax environment	Corporate tax system	- Constructing valid control group	- Constructing valid control group	- Different corporate rate	•••
	State R&D tax credit program	〃	〃	- Different credit rate - Detailed utilization rules	•••
Different business environment	Human capital	- Labor force	- Labor force - Skilled worker		••
	Initial economic condition	- Unemployment rate	- Unemployment rate - Income	- Population	•••
	Dynamic economic condition		- Population change	- Income change, - Employment change	••
	Market size Potential investment capacity	- GSP, PI	- GSP, PI		•••••
	Agglomeration economies		- Establishment density	- Labor pooling, - Industrial mix	••
Different R&D environment	Possibility of spillovers/ Competition		〃	〃	••
	Level of R&D	- Constructing valid control group - Knowledge stock	- Constructing valid control group - Knowledge stock		•••••
	Federal level R&D support	- Federal funding	- Federal funding		•••••
	R&D performance of univ. and research institutions		- University funding		•••
Different government policy environment	Direct funding	- Federal funding	- Federal funding - SBIR, STTR funds		•••••
	Indirect Tax incentives			- Federal level - State level	•
	Other economic development strategies			- Enterprise zone - Science park	•
Different firm characteristic	Potential investment capacity	- Sales	- Sales		•••••
	Firm size	- Number of employees	- Number of employees		•••••
	Establishment status				•

Besides the above three factors, there are some factors possibly causing another possible threat to validity, which is history (i.e. possible third factors include other events occurring concurrently). These are mostly under different government policy environment and more specifically other local economic development strategies and other R&D-supporting program either at the federal level or at the state level.

In general, the different state tax environment could be controlled except the possible difference from different tax rate and tax credit rate. The possible outcome changes from the overall business environment could be eliminated from observavle variables as well as unobservable variables except the highlighted ones. The systematic differences between comparison groups based on different R&D environments also could be successfully reduced through selected evaluation strategies except the possible differences from the different level of competition. While the different govornemnt policy environment provides one of strong threats to validity in this study, the threat from direct funding can be dealt with by adding some available variables into the specification. At last, the most firm characteristics can be considered within the specifications except the establishment status.

In this chapter, I conducted research design for evaluating the effectiveness of R&D tax credit programs based on the above discussions of the applicable methodology, which is the quasi-experimental design. First, the hypothesis is defined as testing a positive program effect of state R&D tax credit programs and testing different program effects by firm size. Second, in the quasi-experimental setting, plausible rival hypotheses are enumerated under five relevant environments to private R&D. These five environments are the state tax environment, the business environment, the R&D environment, the policy environment, and the firm environment. Third, the evaluation strategies are developed under defining dependent variables, defining the experimental group and the control group, and developing evaluation methods. The selected dependent variables are R&D spending and employment and the selected observational levels of analysis are the state and the firm. The selected methods are the difference-in-differences /difference-in-difference-in-differences method, and the matching methods incorporating the DD method. The covariates and matching variables are chosen in order to deal with rival hypotheses. Based on the selected methods, six types of empirical specifications are listed. Then, the datasets to be used are illustrated. Finally, the summary of the selected methods and the selected levels of analysis is made for clarifying. Through re-evaluating rival hypotheses, the possible threats of validity in this study are *history and selection* from agglomeration economies, competition and human capital, which are mostly relevant to decide the level of R&D especially at the regional level.

In the next chapter, I discuss the empirical findings of the effectiveness of state R&D tax credits. These findings include the analysis of the effects on growth of R&D spending, the analysis of the effects on growth of employment, the analysis of the effects on growth of R&D spending by firm size, and the analysis of the effects on growth of employment by firm size.

7.0 ANALYSIS AND RESULTS

In this chapter, I present the empirical analyses and results for the effectiveness of state R&D tax credits. These analyses include (1) an analysis of the effects on increasing *R&D spending*, (2) an analysis of the effects on increasing *employment*, (3) an analysis of the effects on increasing *R&D spending by firm size*, and (4) an analysis of the effects on increasing *employment by firm size*.

For testing the effectiveness of state R&D tax credit programs, this study adopts the quasi-experimental approach, in which the presumed outcomes of states with tax credits are compared to the outcomes of states with no tax credits. The primary presumed outcome is private R&D spending; the secondary presumed outcome is number of employees. I use multiple observations over multiple time periods to evaluate these outcomes. As with social program evaluation, these comparison groups are not equivalent, which means various relevant factors might be involved in outcome changes. I seek to reduce these systematic differences in comparison groups by using the DD/DDD methods with covariates and the matching methods incorporating the DD method. I also conduct these comparisons by two industrial categories (i.e., the all industry and the high-technology industry), two observational categories (i.e., the state level and the firm level), and two values (i.e., level and log values).

The first two analyses tell us whether state R&D tax credits generate positive effects on increasing private R&D spending and employment, in terms of overall effects across industries.

A positive effect means an additional increase in outcomes solely from tax credits, separate from any increase from other relevant factors. I estimate the effects by comparing states with credits to states with no credits under the basic assumption that the outcome growth of states with no credits could capture any possible increase from other relevant factors. I apply the evaluation strategies outlined in the previous chapter for dealing with nonequivalence between the treatment group and the control group under the quasi-experimental setting. The last two analyses tell us if there is any difference in outcome growth generated from state R&D tax credits depending on firm size. Firm size is defined as small, medium or large and is based on the number of employees. The possibility of different effects by firm size is an important and interesting issue for utilizing tax credit programs and understanding the mechanism involved with private R&D activities. I test for differences in outcome growth by separately measuring the effects by firm size and then comparing them. The evaluation strategies are generally the same as the above analyses and these analyses are conducted only at the firm level.

This chapter is organized as follows: under each analysis, I introduce the major findings and describe the detailed results and interpretations from multiple datasets and multiple methods. I describe the average historical outcome changes of the treatment and control groups to provide background for comparison. The graphs do not depict the analyses conducted exactly in the sense that the program enactment years vary by state. I adjusted for the systematic differences in the two comparison groups thorough several program strategies mentioned in the previous chapter. Each section includes a discussion of the detailed estimates from the DD/DDD method and the matching method, and concludes with a summary of the empirical analyses and findings.

7.1 ANALYSIS OF THE EFFECTS OF R&D SPENDING ON GROWTH

In this section I examine the effectiveness of state R&D tax credit programs in terms of increasing private R&D spending. State R&D tax credit programs primarily aim to increase private R&D spending by reducing the marginal cost of private R&D activities (David, Hall and Toole, 2000). I examine program effects on increasing R&D spending by testing whether the experimental group had *an additional increase in R&D spending* relative to the control group before and after the program. The experimental group is defined as states having tax credits and the control group is defined as states not having tax credits. The estimated effects include direct effects for the firms receiving credits and indirect effects for the firms benefiting from credits through spillovers.

For the all industry, I find the average outcome growth of the experimental group before and after the program is *larger* than that of the control group, indicating there is an additional increase of R&D spending generated from the credits. Indeed these positive program effects are statistically significant from both the DD methods and the matching methods, and at both the state level and the firm level.[113] Based on multiple indications of positive effects, I conclude that state R&D tax credits positively affect the growth of R&D spending for the all industry and that this is a reliable and strong conclusion.

For the high-technology industry, I also find that the average outcome growth of the experimental group is *larger* than that of the control group. However, only a few estimates are statistically significant from both evaluation methods. This result suggests that there is a positive

---

[113] The R&D spending data for the all industry is constructed as the aggregate at the state level and as the break-down by 3-digit SIC at the firm level.

program effect on R&D spending for the high-technology industry but this is not strongly supported. Due to data availability, this analysis could only be conducted at the firm level.[114]

In sum, R&D tax credits have positive effects on the growth of R&D spending for both the all industry and the high-technology industry. This tells us that state R&D tax credits have the effect of increasing R&D spending not only for the high-technology industry but also for the all industry. In other words, state R&D tax credits generate the broad effect of increased R&D spending. This evidence supports the appropriateness of this credit on the grounds that the primary purpose of this credit law has been fulfilled. Detailed analyses follow in the next section.

7.1.1 R&D spending for the all industry

For understanding overall changes in R&D spending over time by the treatment and control groups, I first examine the yearly mean distribution of R&D spending by the two groups. Figure 17 shows the historical distribution of the state-level data and Figure 18 shows the historical distribution of the firm-level data, in level and log values. These figures reveal that there are steady increases in R&D spending over time at both observational levels and both values, and these steady increases by the two groups follow similar trends with some variations.

In the case of the state-level data, the yearly means of the treatment group are observed mostly at the higher point than the corresponding value of the control group in both level and log values. This suggests that states with credits had a higher level of private R&D activities in terms of average spending per year.

[114] There is no available data on state level R&D spending with industry break-downs.

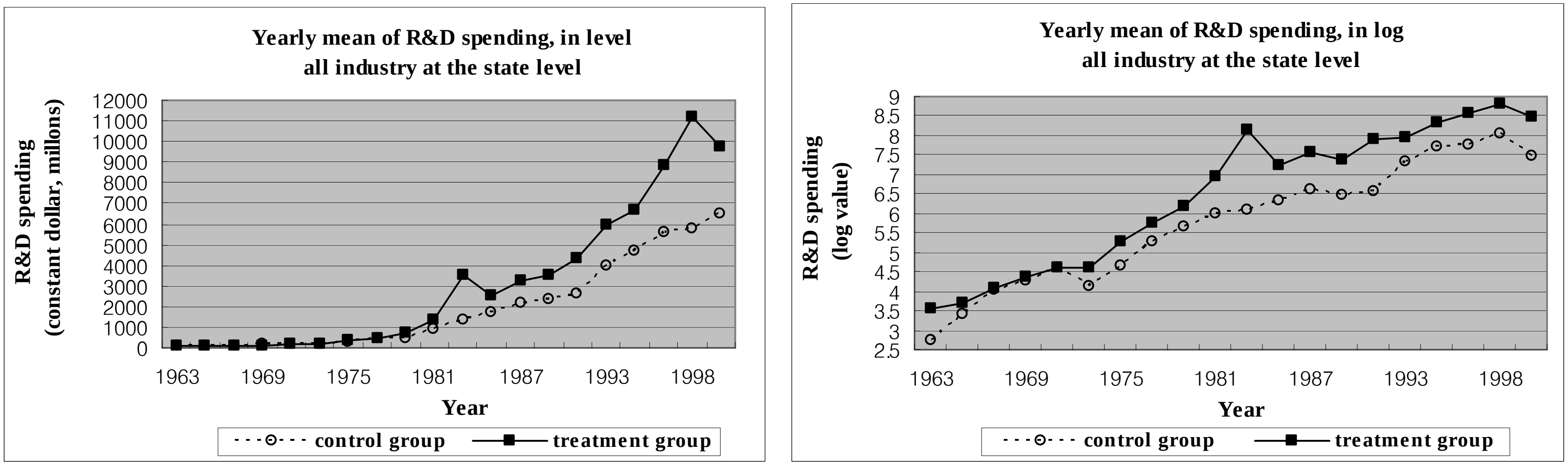


Figure 17. Comparison of historical distribution of R&D spending between the treatment and the control groups for the all industry based on the state-level data

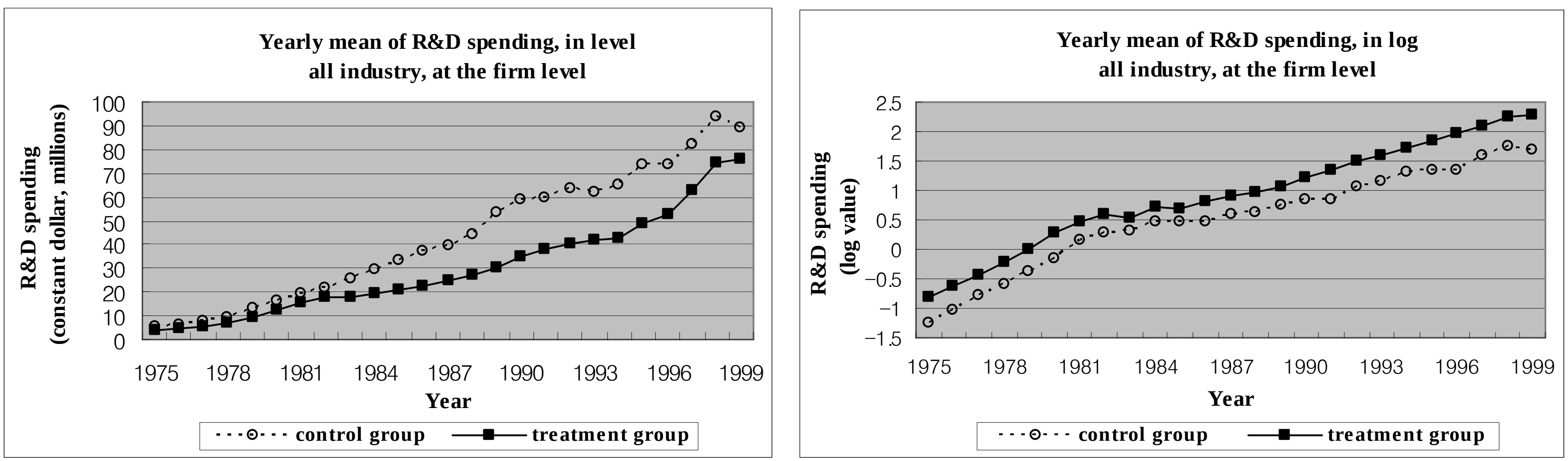


Figure 18. Comparison of historical distribution of R&D spending between the treatment and the control groups for the all industry based on the firm-level data

In the case of the firm-level data, the yearly means of the control group are higher than the corresponding yearly means of the treatment group in level values, but lower in log values. This indicates that the average R&D spending per firm in states with no credits is higher than the R&D spending in states with credits because states with no credits include some observations of high values in R&D spending in level (see Figure 27). These high values create a high mean in level but the effects of these higher values weaken the mean in log because the log transformation reduces the differences of level values.[115] In the empirical specification, the differences in outcome changes shown in Figures 17 and 18 are adjusted by adding covariates and fixed effects for year, state, and industry in the DD methods, or by finding a matched pair for each observation from selected variables in the matching method.

Table 40 shows the detailed estimates of the effects on R&D spending for the all industry by using the DD method. First, most of the DD estimates for program effects are positive at both observational levels (i.e., the state level and the firm level) in both values (i.e., level and log value). At the state level, four possible estimates are obtained from level and log analyses using either the GSP or the PI and other covariates. Among them, both level estimates are positive and statistically significant and both log estimates are positive but only one is significant. At the firm level, both possible estimates (one level estimate and one log estimate) are positive and significant.

Second, average yearly effects of R&D tax credits across states for the all industry can be measured by the 95% confidence intervals of significant estimates. At the state level, the average effects per year range from 250 to 1,000 million 1984 constant dollars and between

[115] For example, 1,000 is transformed by 10 and 100 is transformed by 1, which results in 900 difference in level but 9 difference in log.

0.1% and 24% at the percentage term. At the firm level, the average program effects per year range from 300,000 to 4,000,000 1984 constant dollars and between 2% and 6% at the percentage term. These measured effects include direct effects for firms receiving credits and indirect effects for firms receiving benefits from spillovers.

Table 40. The DD estimates of effects on R&D spending for the all industry

R&D spending for the all industry	Level			Log		
	State		Firm	State		Firm
Program effect (T·Post)	628.9** (189.0)	404.5** (125.4)	1.98** (0.81)	0.10 (0.09)	0.13* (0.07)	0.04** (0.01)
Gross State Product (Lagged)	0.005 (0.003)			0.68 (0.51)		
Personal income (Lagged)		0.002 (0.005)			0.33 (0.50)	
Sales (Lagged)			0.002* (0.001)			0.04** (0.005)
Unemployment rate	-91.64** (38.43)	-63.81* (36.32)	-0.12 (0.07)	-0.09 (0.09)	-0.12 (0.09)	0.020 (0.018)
Labor force	0.26 (0.30)	0.38* (0.20)	0.0015** (0.0007)	-0.71 (0.62)	-0.18 (0.56)	0.009* (0.005)
Knowledge stocks	0.83 (0.51)	0.05 (0.33)	0.002 (0.001)	0.42* (0.25)	0.42 (0.27)	0.005** (0.001)
Private R&D spending (Lagged)	0.81** (0.10)	0.94** (0.08)	0.97** (0.03)	0.43** (0.11)	0.50** (0.08)	0.92** (0.006)
Federal funding (Lagged)	-0.15 (0.10)	-0.10 (0.09)		-0.03 (0.04)	-0.04 (0.03)	
N	185	239	36977	185	239	33564
Adjusted R^2	0.9757	0.9748	0.9687	0.9659	0.9662	0.9396

Notes
· ** Indicates statistical significance at the level of 95%
· * Indicates statistical significance at the level of 90%.
· In the level analysis, private R&D spending and federal funding are measured in constant 1984 million dollars. The GSP, the PI, and the sales are measured in constant 1984 million dollars. Labor forces are measured in thousands. Knowledge stocks are measured as patent stocks depreciated at a rate of 10%, in units of tens. Unemployment rate is measured in percentage.
· In the log analysis, all variables are transformed by natural log.
· The numbers in parentheses indicate heteroskedasticity-consistent and serial-correlation-consistent standard errors.

The analysis using the DD method shows some relationships between R&D spending and relevant factors for the all industry. In the empirical specification, there are six kinds of covariates: lagged GSP/PI/sales, labor forces, unemployment rates, knowledge stocks, lagged private R&D spending, and lagged federal funding. For the firm-level analysis, all variables are constructed at the regional level except lagged sales and private R&D spending, which are constructed at the firm level. The lagged GSP/PI/sales variables reveal positive relationships with private R&D spending in both level and log analyses, and some of them are significant. This indicates that the overall economic investment capacity is important for increasing private R&D spending. At the state level, the unemployment rate reveals significant and negative relationships with private R&D spending, which suggests that the overall economic condition is also important for deciding the level of private R&D spending. At both observational levels, and especially in log analyses, knowledge stocks reveal positive and significant relationships, suggesting that the overall regional level of R&D positively affects private R&D activities. The lagged private R&D spending also positively and significantly increases private R&D spending at both observational levels.[116] The relationship between public funding and private funding is unclear from insignificant estimates. The effect of labor forces is also unclear based on mixed and mostly insignificant estimates, which means that the overall level of human capital is less relevant than the level of private R&D spending for all industry.

[116] The fact that "the coefficients of the lagged dependent variable (which is the lagged private R&D spending) is nearly equal to one" could indicate the possibility of the unit root problem in single time series data analysis, resulting in the spurious regression. However, it is not true in this study because the unit root problem is not serious in fixed-effect panel analysis. In case of single time-series data, the noise is strong but this noise could be substantially reduced by averaging the cross-sectional differences in panel analysis. About more details for possible different interpretations depending on the relationship between dependent variable and independent variables, see Phillips and Moon (1999), Phillips and Moon (2000), and Kao (1999).

Next, the detailed matching estimates of the effects on R&D spending for the all industry are shown in Table 41. The matching method is only applied to the firm-level analysis. There are three kinds of matching estimates, which are nearest matching estimates either using Mahalanobis metrics without caliper (specifying number of matches as 1) or using propensity score metrics with caliper (specifying number of matches as 1 or 3). The significance of the estimates is obtained from the bootstrapping method. All the matching estimates are positive and one out of six possible estimates is significant. The average yearly program effect from the matching methods, measured by the 95% confidence intervals of significant estimates, is from 5 to 18 million 1984 constant dollars per firm in a year.

Table 41. The matching estimates of effects on R&D spending for the all industry at the firm level

R&D spending for the all industry			# of observations		Estimate	Standard error
			Treatment	Control		
Level	Mahalanobis metrics		753	384	11.73**	3.28
	Propensity score metrics	Caliper=0.005 # of matched=1	249	384	6.18	13.66
		Caliper=0.005 # of matched=3	249	384	5.52	8.63
Log	Mahalanobis metrics		753	384	0.04	0.16
	Propensity score metrics	Caliper=0.005 # of matched=1	350	384	0.06	0.14
		Caliper=0.005 # of matched=3	350	384	0.09	0.11

Notes: · Matching is performed based on the nearest neighbor matching method.
· Matching variables are lagged R&D spending, lagged sales, knowledge stocks, labor forces, firm size, per capita income, share of high school graduates, share of college graduates, SBIR funds, university funds, population change, crime rate, poverty rate, unemployment rate, and manufacturing establishment density.
· All dollar values are converted in constant 1984 dollars. R&D spending is measured in millions.
· Standard errors are estimated by using the bootstrapping method.

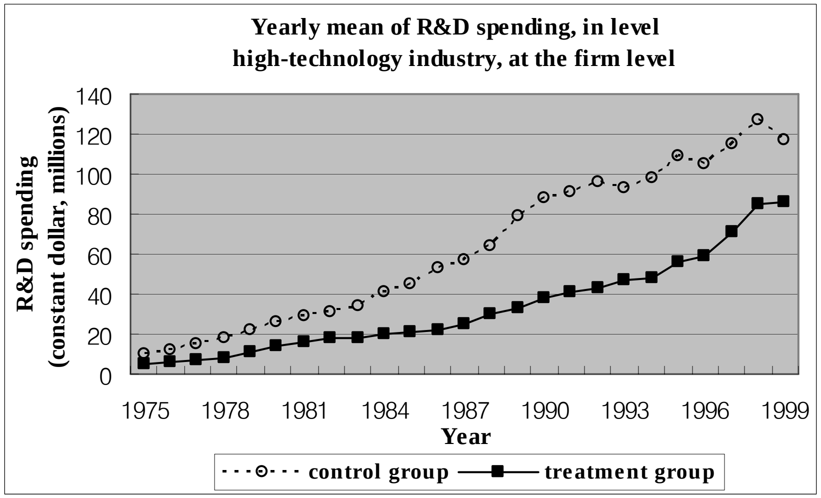
In sum, state R&D tax credit programs generate positive effects on increasing private R&D spending for the all industry based on significant and positive estimates from both observational levels, both analytical methods, and in both level and log values. In the next section I examine the effects on R&D spending for the high-technology industry.

7.1.2 R&D spending for the high-technology industry

For analyzing the effects on R&D spending on the high-technology industry, only the firm-level data is available. First, Figures 19 and 20 show how private R&D spending compares between the treatment and control groups. Figure 19 shows the yearly average of firm R&D spending in the high-technology industry by using level and log values. The average R&D spending per firm in the high-technology industry has increased over time and the historical distribution of the yearly average is mostly similar between states with credits and states without credits. As shown in Figure 18 with the previous analysis for the all industry, average R&D spending per firm in the control group is higher than the firm average R&D spending in the treatment group in level and vice versa in log. This is a result of the fact that the large amount of variation in R&D spending within the control group is relatively adjusted for by transforming the log value.[117]

Figure 20 shows the yearly average of firm R&D spending in the high-technology industry after differencing the yearly average of firm R&D spending in the non-high-technology industry to capture possible outcome change differences between the treatment and control groups by adding the less-influential group. However, the differencing does not reduce the outcome change differences for R&D spending for the high-technology industry. This implies that the non-high-technology industry might not be a valid control group for analyzing private R&D spending for the high-technology industry. Again, the empirical analyses could reduce these differences by adding some covariates and fixed effects for year, state, and industry in the DD/DDD methods and by finding the closest matched pairs in the matching methods.

[117] The existence of firms with substantially large R&D spending is shown in figure 30.



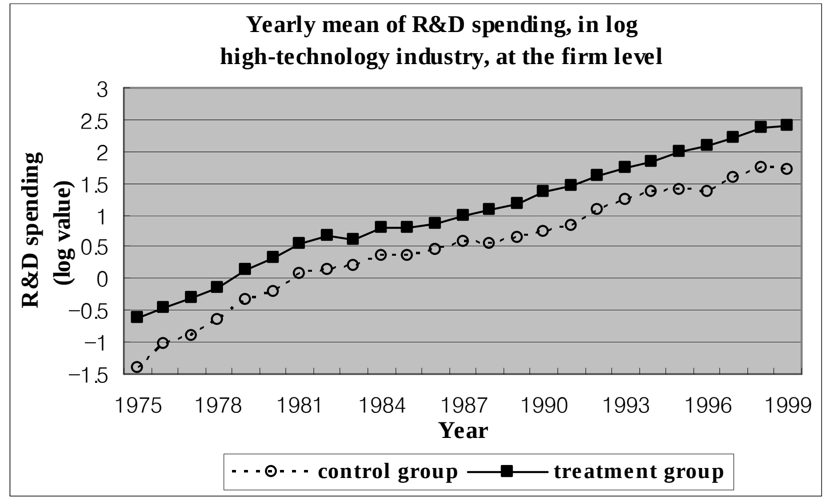
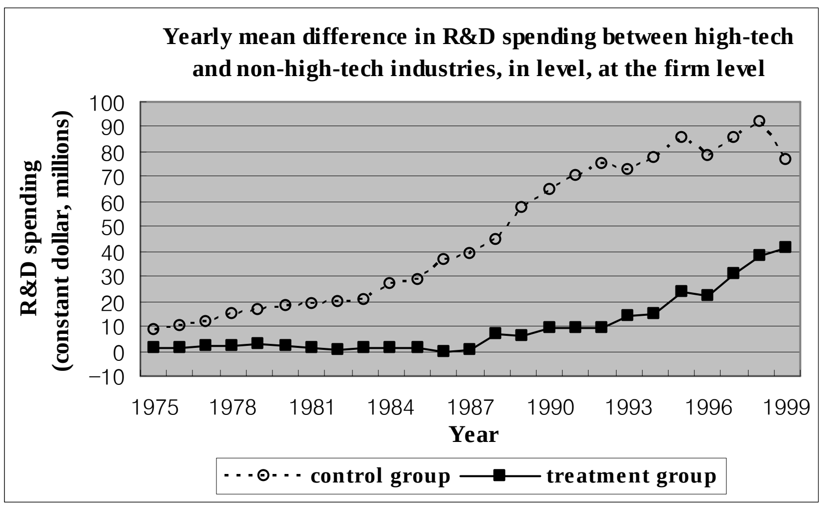


Figure 19. Comparison of historical distribution of R&D spending between the treatment and the control groups for the high-technology industry based on the firm-level data



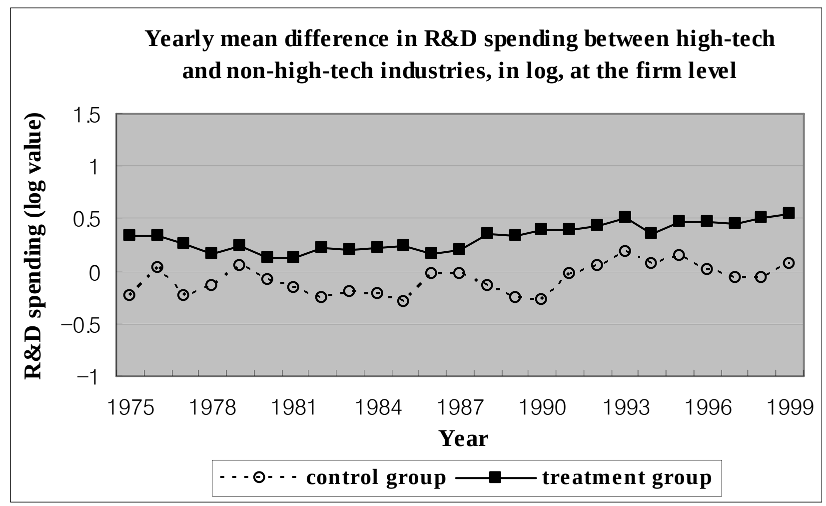


Figure 20. Comparison of historical distribution of R&D spending between the treatment and the control groups for the high-technology industry (after differencing the non-high-technology industry) based on the firm-level data

The detailed DD/DDD estimates of the effects on R&D spending for the high-technology industry are represented in Table 42. The state-level R&D spending data is not broken down by industry. Therefore, I was only able to analyze the high-technology industry data at the firm level. For the analysis of the high-technology industry, two DD methods are applied. One is the DD method that is conducted by comparing outcomes of states having credits with those of states not having credits and using only the high-technology industry data. The other is the DDD method that is conducted by comparing outcomes of states having credits with those of states not having credits, and additionally comparing outcomes for the high-technology industry with those for the non-high-technology industry within each group as the influential group and the less-influential group. Three out of four possible DD/DDD estimates reveal positive program effects, but only one of them is statistically significant.

State R&D tax credit programs have an effect of increasing R&D spending by approximately 200,000 ~ 5,500,000 1984 constant dollars per high-technology firm per year based on 95% confidence intervals.

In the analysis using the DD/DDD method, I also find some relationships between R&D spending and relevant factors for the high-technology industry. There are five kinds of covariates: lagged sales, labor forces, unemployment rates, knowledge stocks, and lagged private R&D spending. Labor forces, the unemployment rate, and knowledge stocks are measured at the regional level while sales and private R&D spending are measured at the firm level. The coefficients for lagged sales, knowledge stocks, and lagged private R&D spending are significant and positive in both level and log analyses. Based on these findings, the firm's overall investment capacity, the region's level of R&D, and the firm's previous R&D activities are important influential factors for increasing current private R&D spending. These positive relationships are identically revealed at the all industry analysis as well, which suggests that the influences of overall environments on increased R&D spending are similar for all and high-technology industries.

Table 42. The DD/DDD estimates of effects on R&D spending for the high-technology industry

R&D spending for the high-technology industry	Level		Log	
	DD	DDD	DD	DDD
T.post	2.85**	1.90	0.02	-0.03
	(1.34)	(1.27)	(0.02)	(0.02)
Sales (Lagged)	0.002*	0.002*	0.03**	0.04**
	(0.001)	(0.001)	(0.006)	(0.005)
Unemployment rate	-0.15	-0.12	0.02	0.01
	(0.12)	(0.07)	(0.02)	(0.01)
Labor force	0.0012*	0.0015**	0.019**	0.008
	(0.0007)	(0.0007)	(0.006)	(0.005)
Knowledge stocks	0.004**	0.002	0.009**	0.005**
	(0.002)	(0.001)	(0.002)	(0.001)
Private R&D spending (Lagged)	0.96**	0.97**	0.93**	0.92**
	(0.03)	(0.03)	(0.007)	(0.006)
N	22373	36977	20557	33564
Adjusted R^2	0.9692	0.9687	0.9384	0.9396

Notes:
· ** Indicates statistical significance at the level of 95%
· * Indicates statistical significance at the level of 90%.
· In the level analysis, private R&D spending is measured in constant 1984 million dollars. The sales are measured in constant 1984 million dollars. Labor forces are measured in thousands. Knowledge stocks are measured as patent stocks depreciated at a rate of 10%, in units of tens. Unemployment rate is measured in percentage.
· In the log analysis, all variables are transformed by natural log.
· The numbers in parentheses indicate heteroskedasticity-consistent and serial-correlation-consistent standard errors.

Next, all possible matching estimates for program effects on R&D spending for the high-technology industry show positive effects, and one is statistically significant as shown in Table 43. The average yearly program effect for high-technology firms, which is measured by the 95%

confidence interval of the significance estimate, ranges from 9 to 27 million 1984 constant dollars.

Table 43. The matching estimates of effects on R&D spending for the high-technology industry at the firm level

R&D spending for the high-technology industry			# of observations		Estimate	Standard error
			Treatment	Control		
Level	Mahalanobis metrics		509	216	18.40**	4.38
	Propensity score metrics	Caliper=0.005 # of matched=1	167	216	13.13	18.74
		Caliper=0.005 # of matched=3	167	216	17.58	16.41
Log	Mahalanobis metrics		509	216	0.43	0.32
	Propensity score metrics	Caliper=0.005 # of matched=1	108	216	0.06	0.25
		Caliper=0.005 # of matched=3	108	216	0.05	0.25

Notes
· Matching is performed based on the nearest neighbor matching method.
· Matching variables are lagged R&D spending, lagged sales, knowledge stocks, labor forces, firm size, per capita income, share of high school graduates, share of college graduates, SBIR funds, university funds, population change, crime rate, poverty rate, unemployment rate, and manufacturing establishment density.
· All dollar values are converted in constant 1984 dollars. R&D spending is measured in millions.
· Standard errors are estimated by using the bootstrapping method.

In sum, state R&D tax credit programs generate the increase in private R&D spending for the high-technology industry based on significant and positive estimates from both analytical methods, in level analyses. However, it is not strong conclusion based on few significant estimates.

In this section, I examined the effects of state R&D tax credit programs on R&D spending. Based on the significant estimates from the DD/DDD methods and the matching methods, I concluded that state R&D tax credit programs have the effect of increasing private R&D spending for both the all industry and the high-technology industry. In the next section, I examine the effects of state R&D tax credit programs on employment.

7.2 ANALYSIS OF THE EFFECTS OF EMPLOYMENT ON GROWTH

In this section I discuss the effectiveness of state R&D tax credit programs by assessing the effects on overall employment growth across industries. The hypothesis to be tested is whether there is an additional increase in employment for states with tax credits before and after the program relative to states with no tax credits. While state R&D tax credit programs originally intend to increase private R&D spending, increasing employment is also expected from this program as a secondary effect. Namely as one of regional economic development policies, state R&D tax credit programs also intend to increase employment through including wages for R&D-relevant workers in qualified R&D expenditures by construction. Therefore, this study includes the analysis of effects on employment levels in order to ascertain the overall positive effects of state R&D tax credits in the grounds of employment growth, which is a primary concern of nearly every regional economic development strategy.

For the all industry, there is no statistically significant and positive estimate from both analytical methods. It tells us that state R&D tax credits do not affect to increase employment for the all industry. For the high-technology industry, there are some significant and positive estimates from both analytical methods, only in level analyses. It tells us that state R&D tax credits positively affect to increase employment for the high-technology industry but this is weak evidence based on few significant estimates.

Therefore, I conclude that there are positive program effects on the employment growth for the high-technology industry, while there is no effect for the all industry. It suggests that the effects of state R&D tax credits on employment growth are not strong enough to spread out for the all industry and instead, are limited to the high-technology industry. Indeed, the effect on employment is not as strong as the effect on R&D spending. The detailed estimates for employment follow in the next section.

7.2.1 Employment for the all industry

For understanding overall changes in employment over time by states with credits and states with no credits, I first compare the historical distribution of yearly mean of number of employees by two groups of states. Figures 21 and 22 represent the historical mean changes in employment of the treatment and control groups for the all industry, at the state and the firm levels, and in level and log values respectively. These historical changes mostly show the similar patterns between two comparison groups, implying that the selected control group can be a valid comparison group.

The state-level data shows steady increases in employment over time, while the firm-level data shows steady decreases. One possible explanation of steady decreases in the average number of employees per firm is that the recent emergence of small-sized firms lowers the average number of employees for recent year observations. Again, at the firm level, the yearly means of states not having credits in level are constantly higher than those of state having credits, while the corresponding values in log are constantly lower than those of states having credits. It might result from the heterogeneity across firms in state not having tax credits, which means the substantially large-sized firms in the control group substantially affect to mean values in level but marginally affect mean values in log.[118]

[118] It is shown in Figure 36.

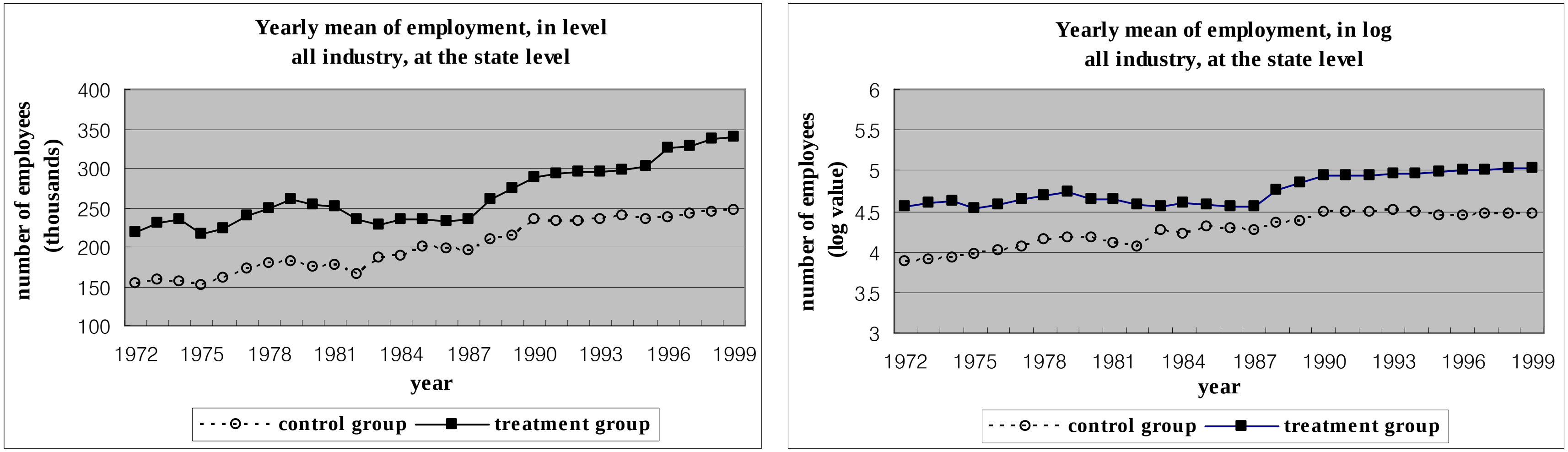


Figure 21. Comparison of historical distribution of employment between the treatment and the control groups for the all industry based on the state-level data

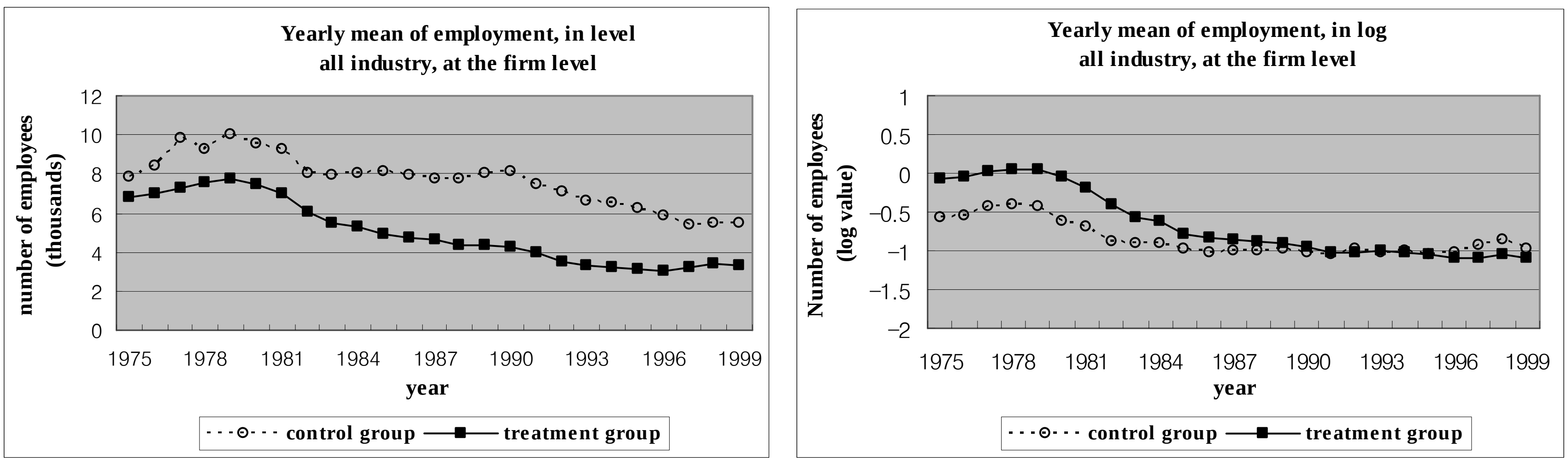


Figure 22. Comparison of historical distribution of employment between the treatment and the control groups for the all industry based on the firm-level data

Table 44 contains the detailed estimates of the effects on employment for the all industry by using the DD/DDD methods. First, all possible DD estimates are positive but none of them is significant. At the state level, there are four possible estimates which are same as the estimates for the analysis of R&D spending. At the firm level, there are two possible estimates, one in level and one in log. All these six estimates are insignificant. Therefore, it tells us that there is no evidence of positive effects on employment growth for the all industry. According to this finding, namely no significant estimates even in a marginal level, the average effect sizes for employment for the all industry cannot be assessed.

Within the analysis using the DD methods, some relationships between the all industry employment and relevant factors are revealed. Under this analysis, six kinds of covariates are included and the unemployment rate is excluded due to the possibility of endogeneity. At the state level, the lagged employment and labor forces show positive and significant estimates in the log analyses. At the firm level, the lagged sales, the lagged private R&D spending, and the lagged employment show the positive and significant estimates while knowledge stocks show the significant and negative estimate mostly from the log analyses. In both observational levels, the previous employment level plays an important role in increasing the current employment level.

Table 44. The DD estimates of effects on employment for the all industry

Employment for all industry	Level			Log		
	State		Firm	State		Firm
T.post	2.36 (6.10)	4.93 (4.61)	0.20 (0.12)	0.02 (0.04)	0.03 (0.02)	0.007 (0.010)
Gross State Product (lagged)	0.007 (0.1)			-0.25 (0.18)		
Personal income (lagged)		0.002 (0.008)			-0.37 (0.24)	
Sales (lagged)			-0.0005 (0.001)			0.02** (0.004)
Labor force	0.005 (0.006)	0.003 (0.003)	0.006 (0.011)	0.36 (0.23)	0.35* (0.18)	0.003 (0.003)
Knowledge stocks	-0.004 (0.01)	0.008 (0.01)	-0.0004 (0.001)	0.03 (0.09)	0.09 (0.07)	-0.002* (0.001)
Private R&D spending (lagged)	0.0006 (0.001)	-0.0009 (0.007)	-0.001 (0.002)	-0.008 (0.03)	-0.03 (0.03)	0.01** (0.002)
Federal funding (lagged)	0.001 (0.002)	-0.0001 (0.001)		-0.003 (0.004)	-0.004 (0.004)	
Employment (lagged)	1.009** (0.007)	1.007** (0.008)	0.99** (0.02)	0.98** (0.01)	0.98* (0.01)	0.93** (0.005)
N	355	450	35435	355	450	32438
Adjusted R^2	0.9971	0.9963	0.9854	0.9931	0.9939	0.9777

Notes:
· ** Indicates statistical significance at the level of 95%
· * Indicates statistical significance at the level of 90%.
· In the level analysis, private R&D spending and federal funding are measured in constant 1984 million dollars. Employment is measured by thousands. The GSP, the PI, and the sales are measured in constant 1984 million dollars. Labor force is measured by millions. Knowledge stocks are measured as patent stocks depreciated at a rate of 10%, in units of tens.
· In the log analysis, all variables are transformed by natural log.
· The numbers in parentheses indicate heteroskedasticity-consistent and serial-correlation-consistent standard errors.

Table 45 contains the detailed estimates for the effects on employment growth for the all industry by using the matching method. The matching method is only applied to the firm-level analysis. The significance of the estimates is obtained from the bootstrapping method. The coefficients of matching estimates in level have mixed signs and none of them is significant. All the coefficients of matching estimates in log are positive but insignificant. Therefore, it suggests that the effects of state R&D tax credit programs on employment growth for the all industry are unclear, as like the results from the DD/DDD estimates. The average effect sizes cannot be obtained due to no significant estimate even in a marginal level.

Table 45. The matching estimates of effects on employment for the all industry at the firm level

Employment for the all industry			# of observations: Treatment	# of observations: Control	Estimate	Standard Error
Level	Mahalanobis metrics		778	419	0.74	0.53
	Propensity score metrics	Caliper=0.005 # of matched=1	383	419	-1.02	1.64
		Caliper=0.005 # of matched=3	383	419	-0.76	1.67
Log	Mahalanobis metrics		778	419	0.04	0.09
	Propensity score metrics	Caliper=0.005 # of matched=1	383	419	0.01	0.15
		Caliper=0.005 # of matched=3	383	419	0.06	0.16

Notes
· Matching is performed based on the nearest neighbor matching method.
· Matching variables are lagged R&D spending, lagged sales, knowledge stocks, labor forces, firm size, per capita income, share of high school graduates, share of college graduates, SBIR funds, university funds, population change, crime rate, poverty rate, unemployment rate, and manufacturing establishment density.
· All dollar values are converted in constant 1984 dollars.
· Standard errors are estimated by using the bootstrapping method.

In sum, the estimates for finding an additional increase in employment in the treatment group relative to the control group before and after the program for the all industry, reveal either positive or negative, and none of estimate is significant from both analytical methods and both observational levels. Accordingly, these results indicate that there is no program effect on employment growth for the all industry. In the next section, I discuss the effects on increasing employment for the high-technology industry.

7.2.2 Employment for the high-technology industry

In this section I discuss program effects on employment for the high-technology industry. The estimates for measuring program effects include the DD and DDD estimates at both observational levels and the matching estimates at the firm level. Figures 23 through 26 compare the differences of yearly mean of employment between the treatment and the control groups for the high-technology industry. Figures 23 and 24 represent the comparisons made of the state-level data and Figures 25 and 26 represent the comparisons made of the firm-level data.

At the state level, the yearly means of high-technology industry employment are not quite varied over time. After differencing the yearly means of employment for the non-high-technology industry, the level values turn out to be negative, which means the employment in the non-high-technology industry is larger than the employment in the high-technology industry. In addition, the yearly values in the control group are higher than those in the treatment group, indicating that the non-high-technology in states with credits are much larger than the non-high-technology in states with no credits. The historical distributions of yearly means of high-technology industry employment between states having credits and states not having credits do not follow a similar pattern and their distributions after differencing the yearly means of employment in the non-high-technology industry do not follow a similar pattern as well.

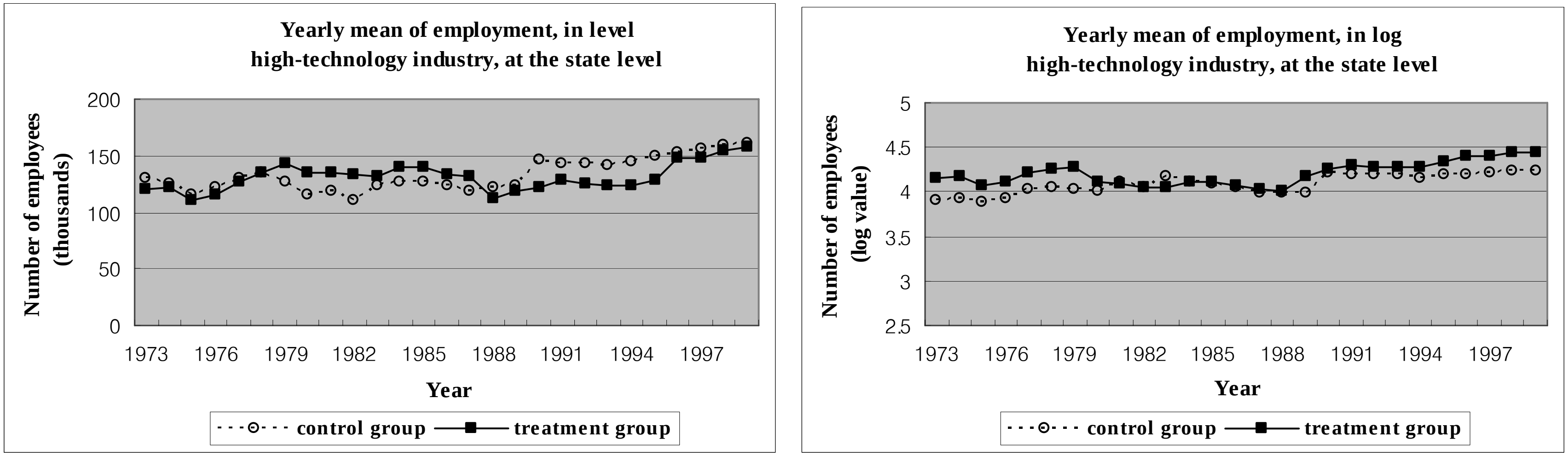


Figure 23. Comparison of historical distribution of employment between the treatment and the control groups for the high-technology industry based on the state-level data

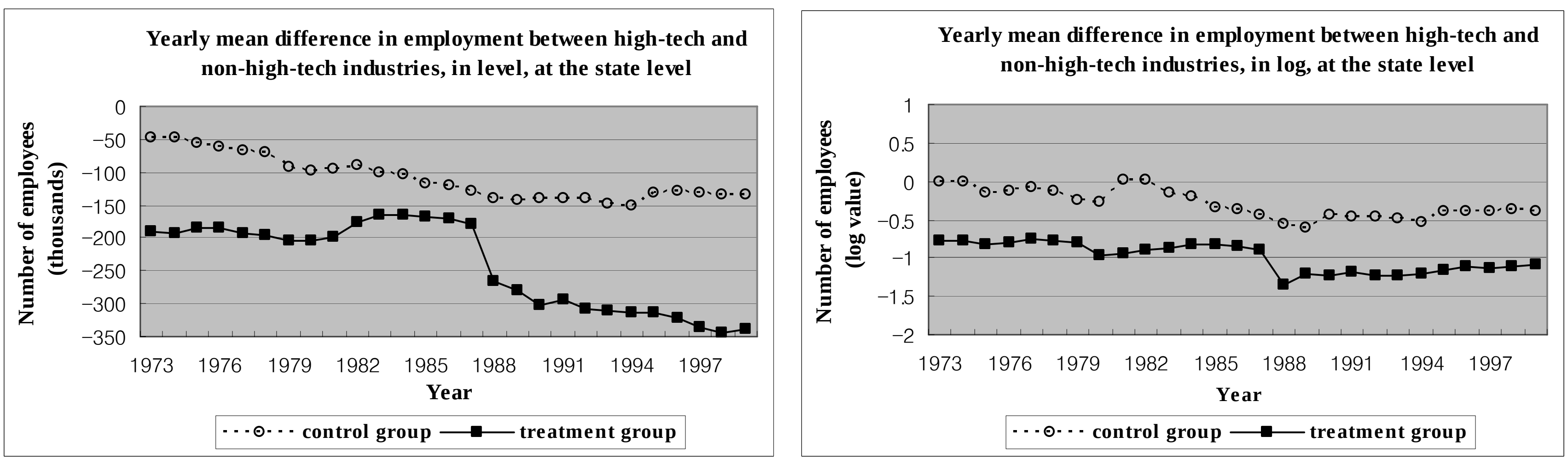


Figure 24. Comparison of historical distribution of employment between the treatment and the control groups for the high-technology industry (after differencing the non-high-tech industry) based on the state-level data

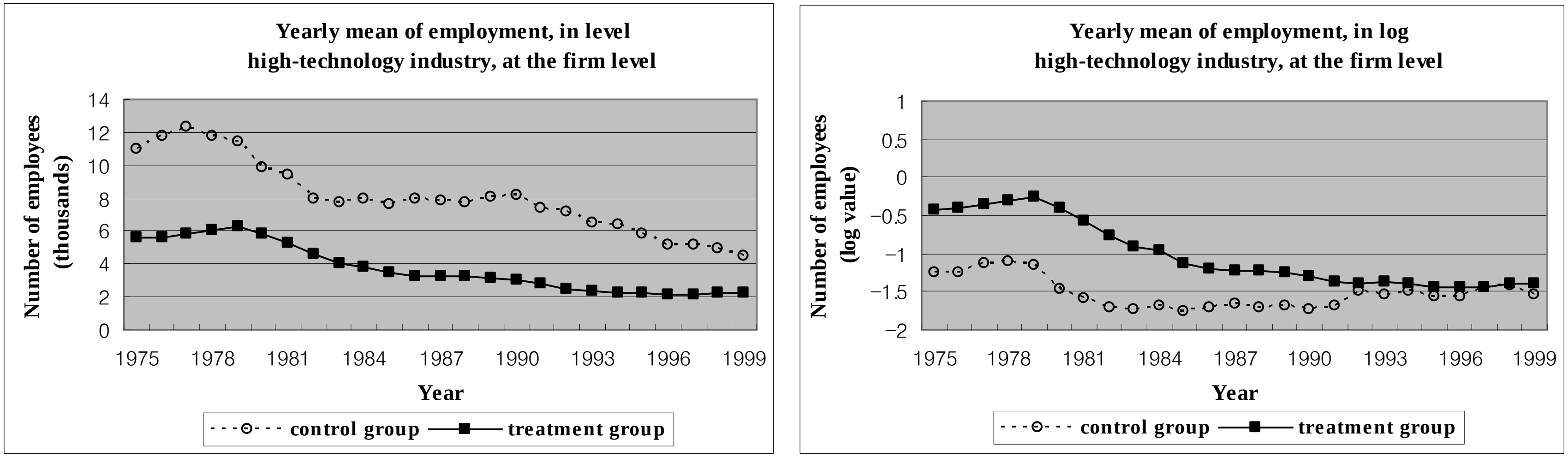


Figure 25. Comparison of historical distribution of employment between the treatment and the control groups for the high-technology industry based on the firm-level data

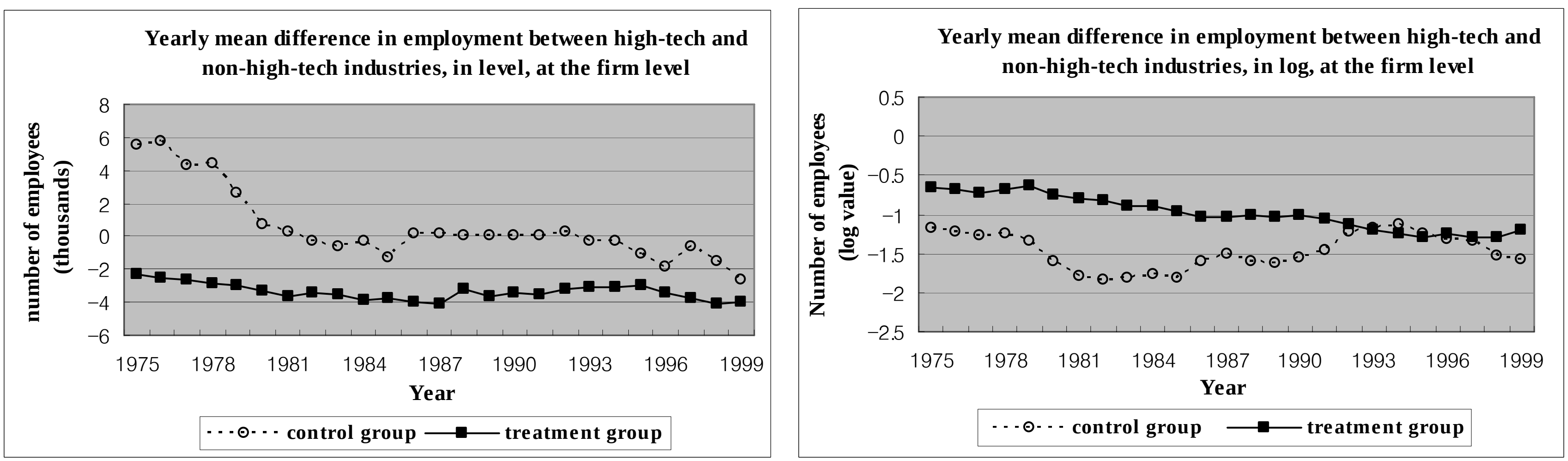


Figure 26. Comparison of historical distribution of employment between the treatment and the control groups for the high-technology industry (after differencing the non-high-technology industry) based on the firm-level data

At the firm level, the historical movements of yearly average employment in the high-technology industry between the treatment and the control groups are similar except a large gap in the early periods. Due to the existence of substantially large-sized firms in the control group,[119] the yearly averages in level are higher in the control group, while the yearly averages in log are higher in the treatment group. These differences of yearly averages in level and log between the treatment and the control groups are also shown in previous analyses for the all industry employment as well. After differencing the yearly means of employment in the non-high-technology industry, the differences between the treatment and the control groups cannot be reduced, which implies that there is a possibility of that the DDD estimates are not the better estimates than the DD estimates. The empirical analyses could reduce these differences by adding some covariates and fixed effects for year, state, and industry in the DD/DDD methods and by finding the closest matched pairs in the matching methods.

The detailed DD/DDD estimates for the high-technology industry employment are represented in Table 46. For the analysis of the high-technology industry, two DD methods are applied. One is the DD method that is conducted by comparing outcomes of states having credits with those of states not having credits and using only the high-technology industry data. The other is the DDD method that is conducted by comparing outcomes of states having credits with those of states not having credits, and additionally comparing outcomes for the high-technology industry with those for the non-high-technology industry within each group as the influential group and the less-influential group. Then, I use the GSP and the personal income as the alternative covariate for the state-level analysis. Therefore, there are six possible estimates in each log and level analysis. The estimates are somewhat mixed, which means some are positive

---

[119] It is shown in Figure 39.

and some are negative. Only one estimate from the analysis at the firm level, by using the DD method, and taken in level value, is marginally significant and positive. These results tell us that there is an only weak evidence of positive program effects on increasing employment for the high-technology industry by measuring an additional increase in employment of the treatment group before and after the program relative to the control group.

Due to no significant estimates at the significance level of 95%, the yearly average program effect, measured by 95% confidence intervals of significant estimates, cannot be assessed for the high-technology industry employment. Instead, based on 90% confidence intervals, state R&D tax credit programs have an effect of increasing approximately 14 ~ 245 employees per high-technology firm in a year. The effect at the state level cannot be assessed in the significance level of 90% as well.

The analyses by using the DD/DDD methods include empirical findings for some relationships between employment and relevant factors for the high-technology industry. At the state level, there are positive relationships with labor forces and the lagged employment, and at the firm level, there are positive relationships with the lagged sales, the lagged firm's R&D spending, labor forces, and the lagged firm's employment. In addition, there are negative relationships with the lagged GSP, knowledge stocks, and federal funding. It tells us that the employment level of the high-technology industry within a state is affected by the overall regional level of human capital and the previous employment level within a state. In addition, it also tells us that the employment level of the high-technology industry for a firm is affected by the overall level of investment capacity, R&D spending, and the previous employment for a firm and the overall regional level of human capital.

Table 46. The DD/DDD estimates of effects on employment for the high-technology industry

Employment for high-technology industry	Level						Log					
	State				Firm		State				Firm	
	DD		DDD		DD	DDD	DD		DDD		DD	DDD
T.post	-2.70 (4.51)	1.06 (3.58)	-9.59 (7.06)	-5.82 (7.28)	0.13* (0.07)	-0.012 (0.15)	-0.03 (0.05)	0.001 (0.03)	-0.05 (0.05)	-0.02 (0.03)	0.01 (0.01)	0.01 (0.01)
Gross State Product (lagged)	-0.001 (0.001)		-0.0003 (0.01)				-0.36 (0.35)		-0.25 (0.20)			
Personal income (lagged)		-0.001 (0.001)		-0.0006 (0.01)				-0.69 (0.60)		-0.38 (0.25)		
Sales (lagged)					-0.0001 (0.0001)	-0.0005 (0.0001)					0.02** (0.005)	0.02** (0.004)
Private R&D spending (lagged)	0.0005 (0.001)	-0.0003 (0.0005)	0.0006 (0.001)	0.0005 (0.007)	-0.003 (0.002)	-0.001 (0.002)	0.05 (0.06)	0.02 (0.04)	-0.002 (0.03)	-0.03 (0.04)	0.01** (0.004)	0.01** (0.002)
Labor force	0.0001 (0.0001)	0.001** (0.0005)	0.006 (0.005)	0.005 (0.003)	0.0001 (0.001)	0.007 (0.01)	0.78 (0.60)	0.76 (0.59)	0.41* (0.24)	0.39* (0.21)	0.01** (0.004)	0.003 (0.003)
Knowledge stocks	-0.001 (0.002)	-0.009 (0.01)	-0.004 (0.02)	0.008 (0.01)	0.0001 (0.0001)	-0.0003 (0.001)	-0.06 (0.11)	0.06 (0.12)	0.03 (0.09)	0.09 (0.07)	-0.001* (0.001)	-0.002* (0.001)
Federal funding (lagged)	-0.003 (0.002)	-0.04** (0.01)	0.001 (0.002)	-0.0009 (0.01)			-0.005 (0.006)	-0.006 (0.01)	-0.003 (0.005)	-0.003 (0.005)		
Employment (lagged)	0.97** (0.03)	0.97** (0.02)	1.004** (0.009)	1.001** (0.01)	1.02** (0.02)	0.99** (0.02)	0.94** (0.05)	0.93** (0.05)	0.97** (0.02)	0.97** (0.02)	0.94** (0.007)	0.93** (0.005)
N	164	206	355	450	21006	35435	164	206	355	450	19444	32438
Adjusted R^2	0.9935	0.9926	0.9970	0.9963	0.9913	0.9854	0.9885	0.9898	0.9932	0.9939	0.9730	0.9777

Notes: · ** Indicates statistical significance at the level of 95% and * indicates statistical significance at the level of 90%. · In the level analysis, private R&D spending and federal funding are measured in constant 1984 million dollars. Employment is measured in thousands. The GSP, the PI, and the sales are measured in constant 1984 million dollars. Labor force is measured in millions. Knowledge stocks are measured as patent stocks depreciated at a rate of 10%, in units of tens. Unemployment rate is measured in percentage. · In the log analysis, all variables are transformed by natural log. The numbers in parentheses indicate heteroskedasticity-consistent and serial-correlation-consistent standard errors.

Next, the detailed matching estimates of the effects on the high-technology industry employment growths are shown in Table 47. The matching method is only applied to the firm level analysis. There are three kinds of matching estimates, which are nearest matching estimates either using Mahalanobis metrics without caliper (specifying number of matches as 1) or using propensity score metrics with caliper (specifying number of matches as 1 or 3). The significance of the estimates is obtained from the bootstrapping method. All the matching estimates show positive ones and only one estimate in level is marginally significant.

Due to no significant estimates at the significance level of 95%, the yearly average program effect, measured by 95% confidence intervals of significant estimates, cannot be assessed for the high-technology industry employment. Instead, based on 90% confidence intervals, state R&D tax credit programs have an effect of increasing approximately 50 ~ 1,300 employees per high-technology firm in a year.

Table 47. The matching estimates of effects on employment for the high-technology industry at the firm level

Employment for the high-technology industry			# of observations		Estimate	Standard error
			Treatment	Control		
Level	Mahalanobis metrics		516	233	0.71*	0.39
	Propensity score metrics	Caliper=0.005 # of matched=1	167	233	0.82	1.77
		Caliper=0.005 # of matched=3	167	233	1.38	1.46
Log	Mahalanobis metrics		516	233	0.07	0.13
	Propensity score metrics	Caliper=0.005 # of matched=1	152	233	0.04	0.24
		Caliper=0.005 # of matched=3	152	233	0.12	0.18

Notes
· Matching is performed based on the nearest neighbor matching method.
· Matching variables are lagged R&D spending, lagged sales, knowledge stocks, labor forces, firm size, per capita income, share of high school graduates, share of college graduates, SBIR funds, university funds, population change, crime rate, poverty rate, unemployment rate, and manufacturing establishment density.
· All dollar values are converted in constant 1984 dollars.
· Standard errors are estimated by using the bootstrapping method.

Consequently, state R&D tax credits generate the positive effects on increasing employment for the high-technology industry, although it is not strong evidence based on only few marginally significant estimates.

In this section, I examined the effects of state R&D tax credit programs on employment. Based on no significant estimates for program effect, I concluded that state R&D tax credits do not affect to increase the all industry employment. Based on few significant and positive estimates, I concluded that state R&D tax credits positively affect to increase the high-technology employment and this is weak evidence. Accordingly, the revealed effect of state R&D tax credits on employment growth is narrow and limited. In the next section, I examine the different effects of state R&D tax credit programs on R&D spending depending on firm size.

7.3 ANALYSIS OF THE EFFECTS OF R&D SPENDING ON GROWTH BY FIRM SIZE

In this section I analyze the effects of state R&D tax credits in terms of different policy impacts on increasing R&D spending by firm size. The hypothesis to be tested is whether there is any different R&D spending growth from state R&D tax credits by firm size. This hypothesis is tested by estimating the average additional increase in R&D spending for states with tax credits before and after the program relative to states with no tax credits for small, medium and large firms separately and then comparing them. Due to data availability, the analysis is performed at the firm level only.

The effects of R&D tax credits on R&D spending for small firms have been receiving substantial interests to policy makers as well as researchers. For policy makers, encouraging

R&D activities of small firms is one of fundamental tasks for sustainable economic developments especially at the regional level. For researchers, the linkage between R&D activities and firm size is one of ongoing interests and policy impact related with this linkage is another interesting topic.

Across all industries, I find statistically significant and positive estimates for all three kinds of firms and these results are detected from both analytical methods (i.e. the DD method and the matching method). Across high-technology industries, I also find statistically significant and positive estimates for all three kinds of firms and these results are also detected from both analytical methods. Interestingly, these findings suggest that not only large firms but also medium and small firms could receive benefits for increasing their R&D spending from state R&D tax credits. Contrarily the common beliefs of limited program effects to large firms, this study provides the strong evidence of broad program effects to all kinds of firms for increasing their R&D spending. By construction, the measured effects include direct effects and indirect effects through spillovers. As the effects on R&D spending are spread out across industries, those effects are also spread out across different sized firms as well. The detailed analysis follows.

7.3.1 R&D spending by firm size for the all industry

For analyzing the effects on R&D spending by firm size across all industries, the effects are estimated separately by three sized firms and then compared each other. Figures 27 through 29 show how private R&D spending compares between the treatment and the control groups, by large, medium and small firms respectively. First of all, the yearly average R&D spending per firm has increased over time, with similar growth patterns especially in early periods, for three different sized firms. These tendencies are more apparent in log values than in level values. For small and medium firms, there are the larger outcome growths in late periods in the treatment group than the control group, implying the possible outcome increases from program at least partially.

As explained above, the yearly means of R&D spending in level for large firms are higher in the control group than in the treatment group over time, while the corresponding values for medium and small firms are higher in the treatment group than the control group. It indicates that medium and small firms in states having credits have spent relatively more R&D spending in average than in states not having credits, while large firms have spent relatively less.

By separately analyzing large, medium and small firms, the heterogeneity between firms can be reduced and it makes comparison groups more comparable. This similarity provides the ground of the better comparison in the sense that the comparison group should be similar with the treatment group at least in pretest periods for applying the DD/DDD methods. The empirical analyses could reduce possible systematic differences between two comparison groups by adding some covariates and fixed effects for year, state, and industry in the DD/DDD methods and by finding the closest matched pairs in the matching methods.

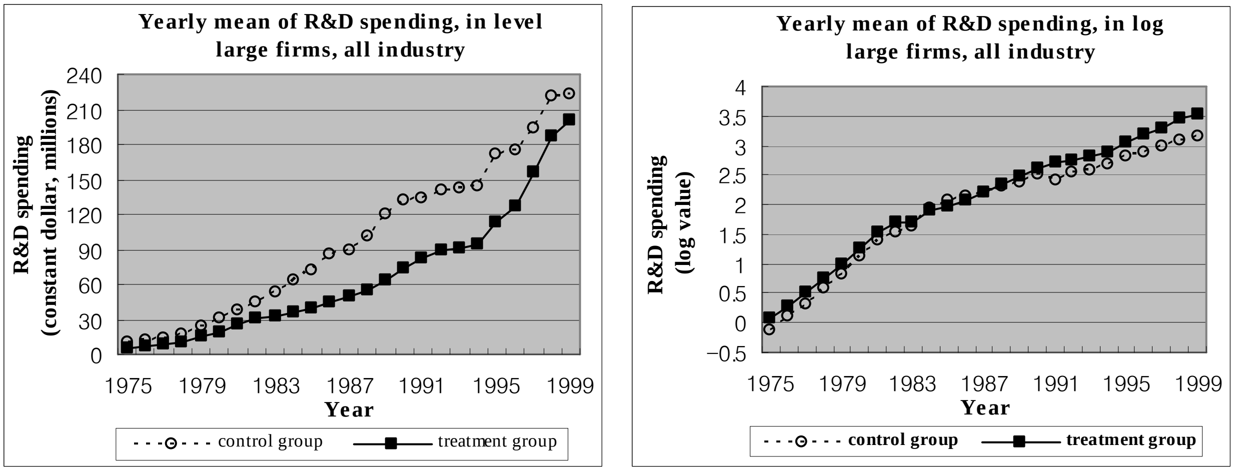


Figure 27. Comparison of historical distribution of R&D spending between the treatment and the control groups for large firms in the all industry

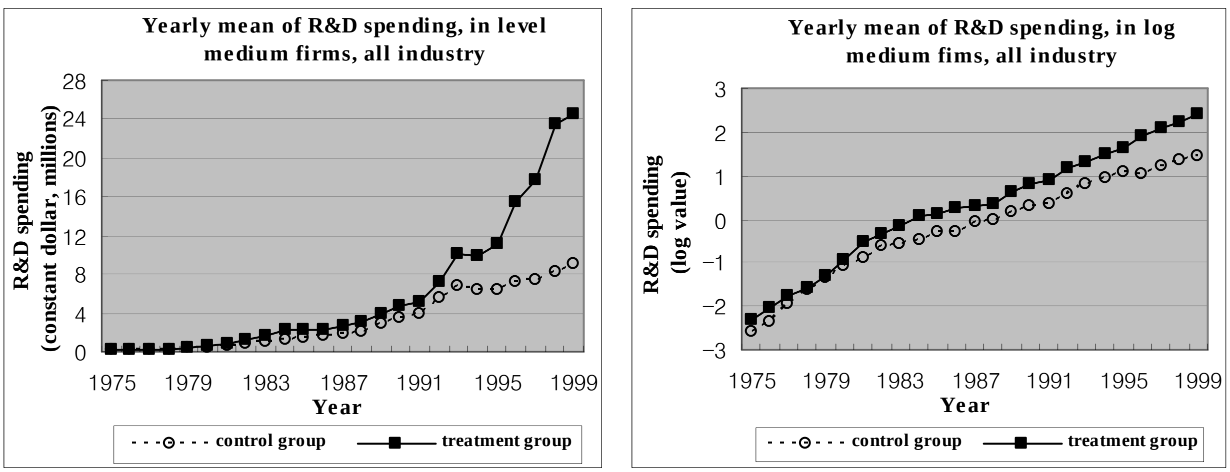


Figure 28. Comparison of historical distribution of R&D spending between the treatment and the control groups for medium firms in the all industry

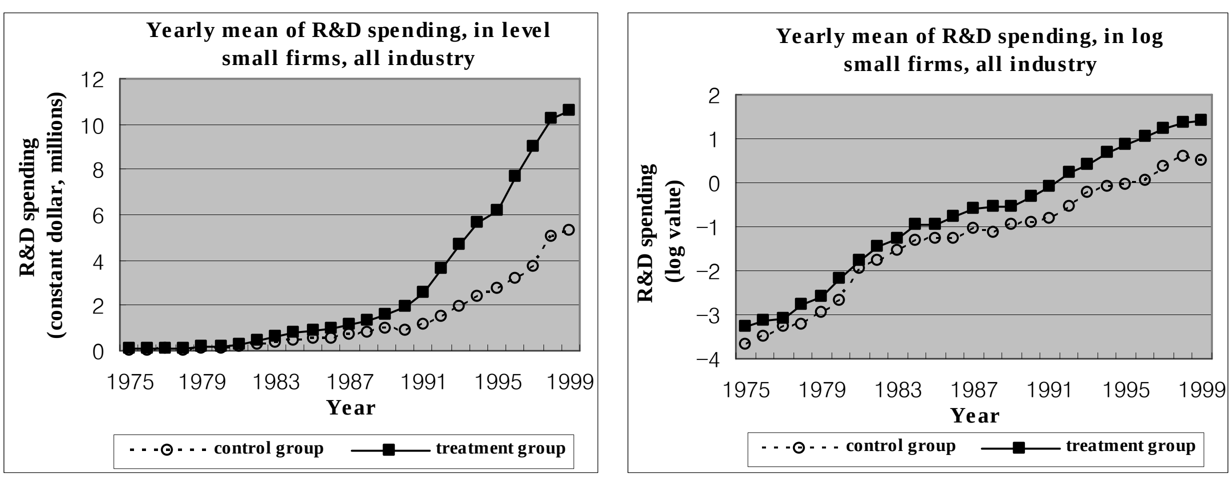


Figure 29. Comparison of historical distribution of R&D spending between the treatment and the control groups for small firms in the all industry

The DD estimates of effects on R&D spending by firm size across all industries are summarized in Table 48. There are two possible estimates, one in level and one in log, in each sized firm. Most of the DD estimates are positive and significant, indicating that there are additional increases in R&D spending in the treatment group before and after the program relative to the control group for all three sized firms. In detail, for medium and large firms, both possible estimates are significant and positive. Meanwhile, for small firms, the level estimate is positive and marginally significant and the log estimate is positive and insignificant. Therefore, the revealed positive effects are weak for small firms but strong for medium and large firms.

Table 48. The DD estimates of effects on R&D spending by firm size across all industries

R&D spending for the all industry	Level			Log		
	Small	Medium	Large	Small	Medium	Large
T.post	0.11* (0.06)	0.95** (0.41)	4.72** (2.03)	0.03 (0.03)	0.08** (0.04)	0.03** (0.01)
Sales (Lagged)	0.006** (0.001)	0.03** (0.01)	0.002** (0.001)	0.02** (0.006)	0.001 (0.01)	0.05** (0.01)
Unemployment rate	0.03** (0.01)	-0.02 (0.05)	-0.17 (0.15)	0.08** (0.04)	-0.01 (0.05)	0.008 (0.01)
Labor force	0.00009 (0.00009)	-0.00007 (0.0003)	0.003 (0.003)	0.03** (0.01)	-0.02 (0.01)	0.006 (0.005)
Knowledge Stocks	0.0003 (0.0002)	0.001** (0.0006)	0.002 (0.003)	0.009** (0.002)	0.003 (0.005)	0.003 (0.002)
Private R&D spending (lagged)	0.98** (0.03)	0.66** (0.16)	0.96** (0.03)	0.85** (0.009)	0.87** (0.01)	0.93** (0.007)
N	12511	5611	18670	10301	4782	17989
Adjusted R^2	0.7809	0.5958	0.9682	0.8400	0.8875	0.9614

Notes:
· ** Indicates statistical significance at the level of 95%
· * Indicates statistical significance at the level of 90%.
· In the level analysis, private R&D spending is measured in constant 1984 million dollars. The sales are measured in constant 1984 million dollars. Labor force is measured in millions. Knowledge stocks are measured as patent stocks depreciated at a rate of 10%, in units of tens. Unemployment rate is measured in percentage.
· In the log analysis, all variables are transformed by natural log.
· The numbers in parentheses indicate heteroskedasticity-consistent and serial-correlation-consistent standard errors.

The average yearly program effects on R&D spending by firm size across all industries are measured by 95% confidence intervals for medium and large firms, and 90% confidence intervals for small firms. For small firms, the average yearly effects range from 9,000 to 200,000 1984 constant dollars. For medium firms, the average yearly effects range from 330,000 to 1,500,000 1984 constant dollars or between 1% and 16%. For large firms, the average yearly effects range from 240,000 to 9,000,000 1984 constant dollars or between 0.3% and 6%. These results show that in terms of the amount of spending, the effects for large firms are larger than those for medium and small firms, while in terms of the % change of R&D spending, the effects for medium firms are larger than those for large firms.

Within the analyses using the DD methods with covariates, I also find some relationships between R&D spending and relevant factors by firm size for the all industry. Among selected covariates, the lagged sales, knowledge stocks, and the previous R&D spending reveal positive relationships with R&D spending for all three kinds of firms. It tells us that the individual firm's overall investment capacity and previous R&D investment, and region's overall R&D environment are important for deciding the current amount of R&D investment regardless of firm size. Interestingly, unemployment rate only significantly affects to small firm's R&D spending with a positive relationship, which implies that small firm's decision for R&D activities do not follow overall economic conditions.

The matching estimates of effects on R&D spending by firm size across all industries are summarized in Table 49. The coefficients of matching estimates are mostly positive. Among them, the estimates in level and by using Mahalanobis metrics are positive and significant for all three sized firms. As like the above DD estimates, the matching estimate for small firms is marginally significant at the significance level of 90% while the matching estimates for medium and large firms are significant at the significance level of 95%. Indeed, the estimates in level and by using propensity score metrics for large firms are also positive and marginally significant, providing strong evidence for positive effects for large firms across all industries.

Based on significant estimates, the average yearly program effects on R&D spending by firm size are measured by either 95% or 90% confidence intervals and in 1984 constant dollars. For small firms, the average yearly effects range from 40,000 to 670,000. For medium firms, average yearly effects range from 10,000 to 3,500,000. For large firms, average yearly effects range from 20,000 to 60,000,000. There is no significant estimate in log, and therefore the average effects in % change cannot be assessed.

Based on the above analysis, I conclude that the state R&D tax credit programs positively affect to increase firm's R&D spending across all industries regardless firm size. However, the revealed effects for small firms are relatively weak and those for medium and large firms are relatively strong based on the marginally significant estimates for small firms and strongly significant estimates for medium and large firms from both analytical methods. In the next section, I analyze the effects on R&D spending by firm size across high-technology industries.

Table 49. The matching estimates of effects on R&D spending by firm size across all industries

R&D spending for the all industry			# of observations		Estimate	Standard error
			Treatment	Control		
Level, small firms	Mahalanobis metrics		252	131	0.36*	0.19
	Propensity score metrics	Caliper=0.005 # of matched=1	38	131	0.20	0.55
		Caliper=0.005 # of matched=3	38	131	0.22	0.55
Level, medium firms	Mahalanobis metrics		110	54	1.76**	0.87
	Propensity score metrics	Caliper=0.005 # of matched=1	14	54	-2.71	4.96
		Caliper=0.005 # of matched=3	14	54	-2.81	4.15
Level, large firms	Mahalanobis metrics		391	199	23.95**	6.02
	Propensity score metrics	Caliper=0.005 # of matched=1	182	199	30.57*	18.37
		Caliper=0.005 # of matched=3	182	199	30.96*	16.62
Log, small firms	Mahalanobis metrics		252	131	0.14	0.16
	Propensity score metrics	Caliper=0.005 # of matched=1	38	131	-0.08	0.51
		Caliper=0.005 # of matched=3	38	131	0.04	0.41
Log, medium firms	Mahalanobis metrics		110	54	-0.05	0.23
	Propensity score metrics	Caliper=0.005 # of matched=1	14	54	0.21	0.54
		Caliper=0.005 # of matched=3	14	54	0.08	0.43
Log, large firms	Mahalanobis metrics		391	199	0.25	0.19
	Propensity score metrics	Caliper=0.005 # of matched=1	118	199	0.03	0.24
		Caliper=0.005 # of matched=3	118	199	0.08	0.23

Notes
· Matching is performed based on the nearest neighbor matching method.
· Matching variables are lagged R&D spending, lagged sales, knowledge stocks, labor forces, firm size, per capita income, share of high school graduates, share of college graduates, SBIR funds, university funds, population change, crime rate, poverty rate, unemployment rate, and manufacturing establishment density.
· All dollar values are converted in constant 1984 dollars. R&D spending is measured in millions.
· Standard errors are estimated by using the bootstrapping method.

7.3.2 R&D spending by firm size for the high-technology industry

In this section I discuss the program effects on R&D spending by firm size across high-technology industries. Figures 30 through 35 show the historical changes in yearly mean of R&D spending by firm size between the treatment and the control groups across high-technology industries in level and log. As like above analyses of effects on R&D spending by firm size across all industries, the historical distribution of yearly means of R&D spending by firm size across high-technology industries follows similar patterns especially in the early periods before 1990. These tendencies are clearer for small and medium sized firms than for large firms. It means that two comparison groups have similar characteristics for increasing their R&D spending in pretest periods, and provide the ground of better comparison. The observable big differences of outcome growth in level for small and medium sized firms in later periods after 1990 (in Figures 32 and 34) also imply the possibility of involvement of other factors with the possible program effects. Again, because of the existence of substantially large sized firms in the control group, the yearly means of R&D spending for large firms are higher in the control group.

Differencing the yearly means of R&D spending of non-high-technology industry for capturing possible outcome changes from other factors except program, which is represented in Figures 31, 33, and 35 by firm size, do not reduce the differences of two comparison groups. Rather, the differences of two comparison groups are similar in level analyses and even larger in log analyses. It implies that the non-high-technology industry group cannot be a valid comparison group as a less-influential group for capturing outcome changes from other relevant factors.

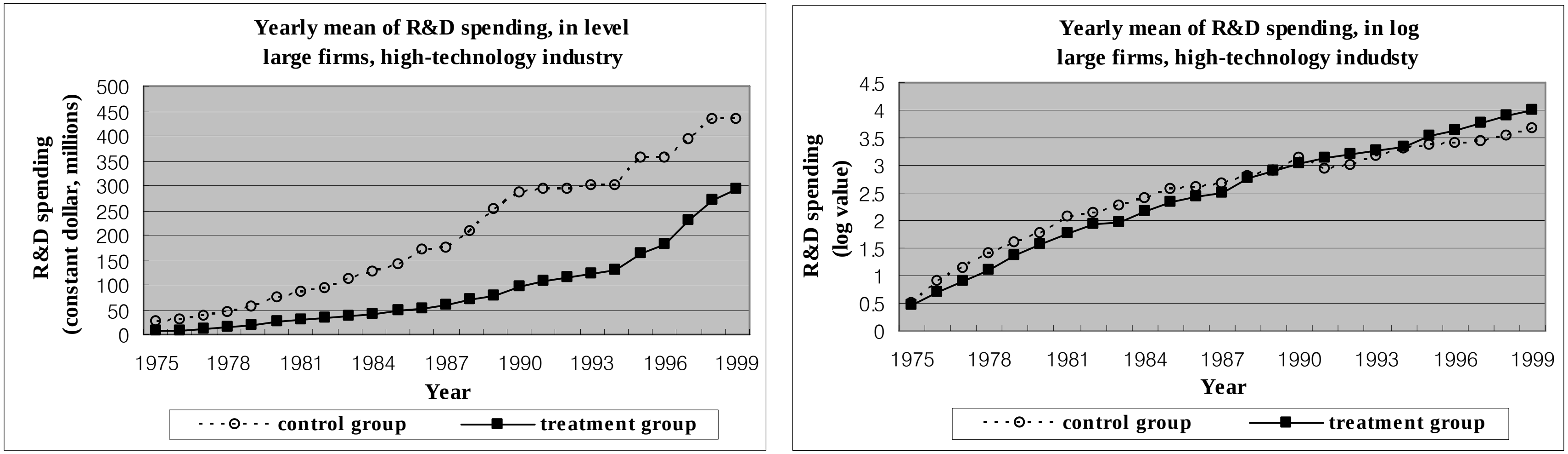


Figure 30. Comparison of historical distribution of R&D spending between the treatment and the control groups for large firms in the high-technology industry

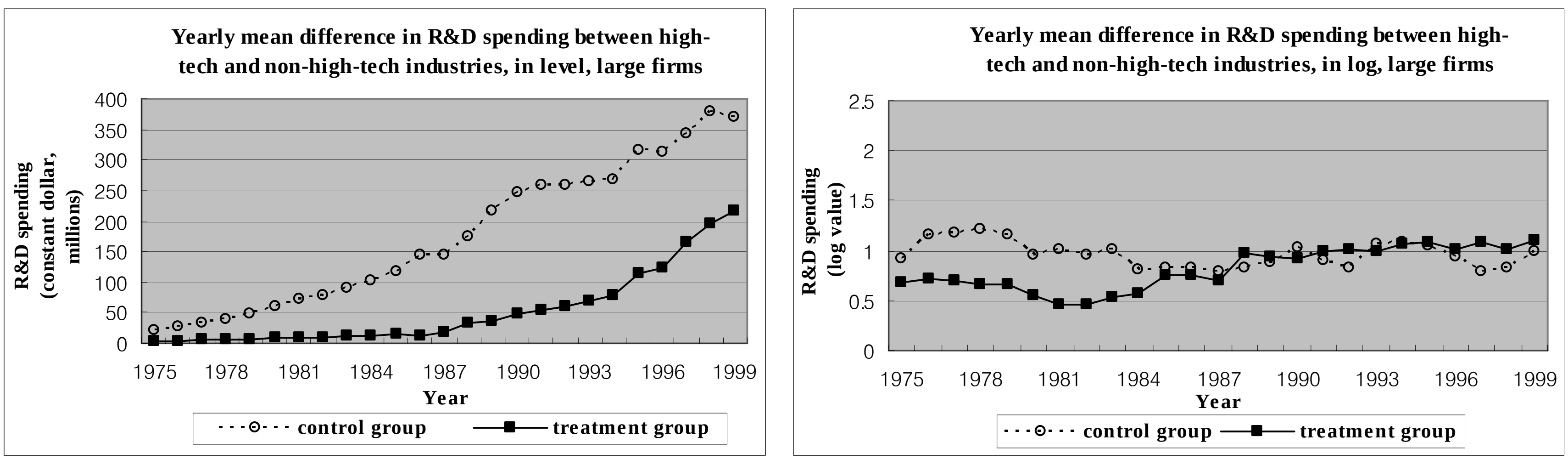
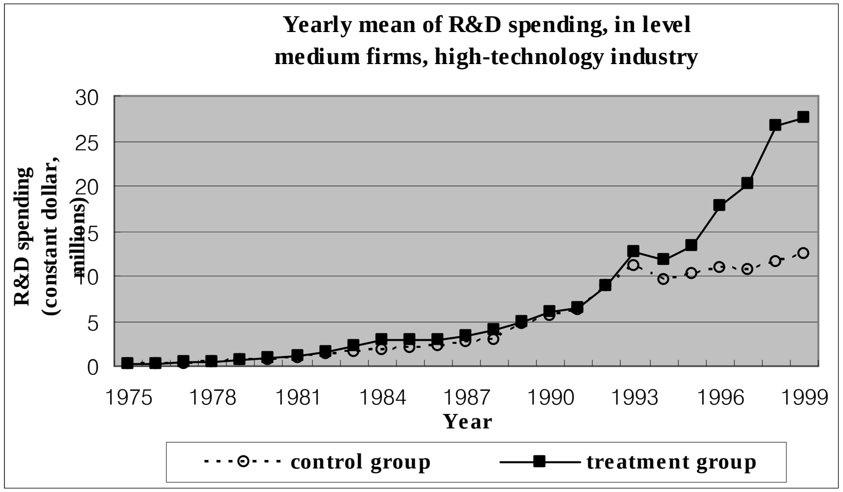


Figure 31. Comparison of historical distribution of R&D spending between the treatment and the control groups for large firms in the high-technology industry (after differencing large firms in the non-high-technology industry)



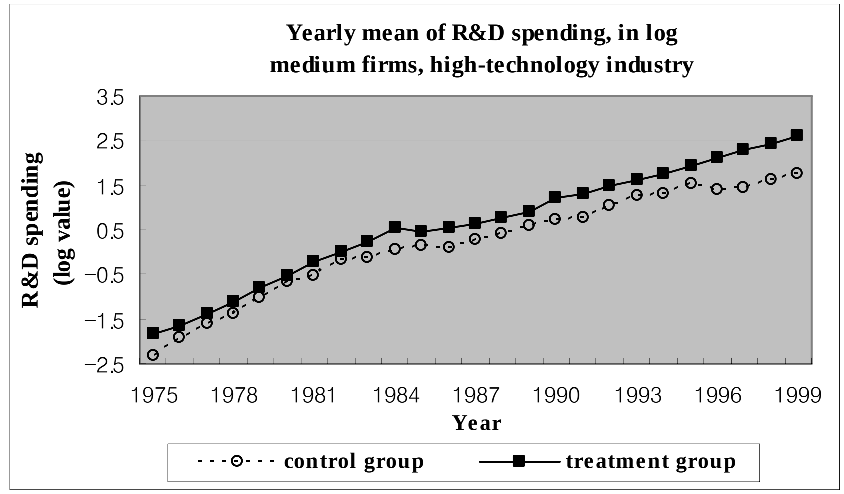
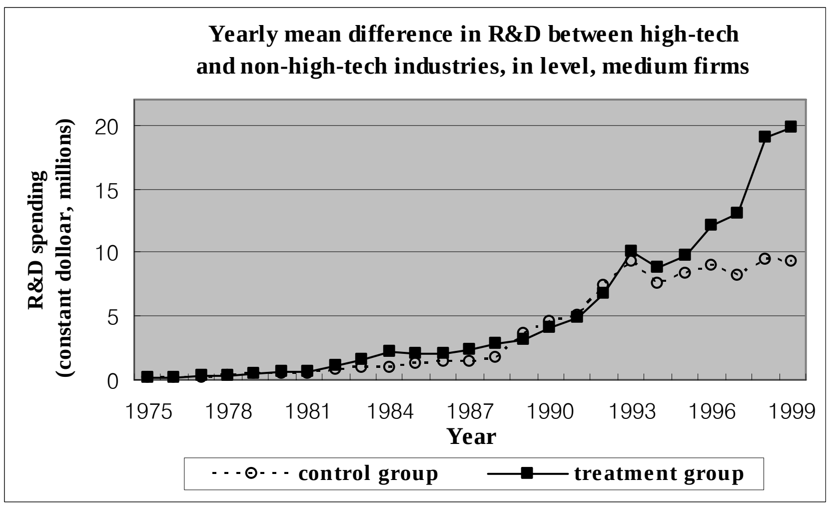


Figure 32. Comparison of historical distribution of R&D spending between the treatment and the control groups for medium firms in the high-technology industry



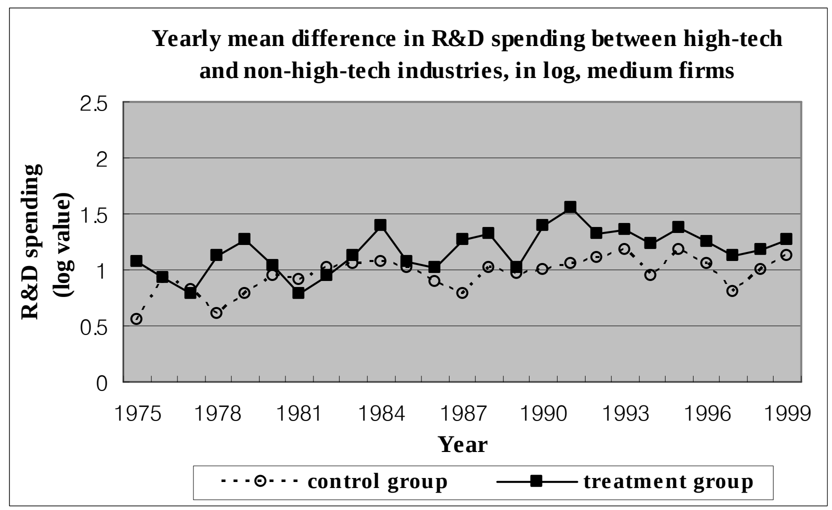


Figure 33. Comparison of historical distribution of R&D spending between the treatment and the control groups for medium firms in the high-technology industry (after differencing medium firms in non-high-technology industry)

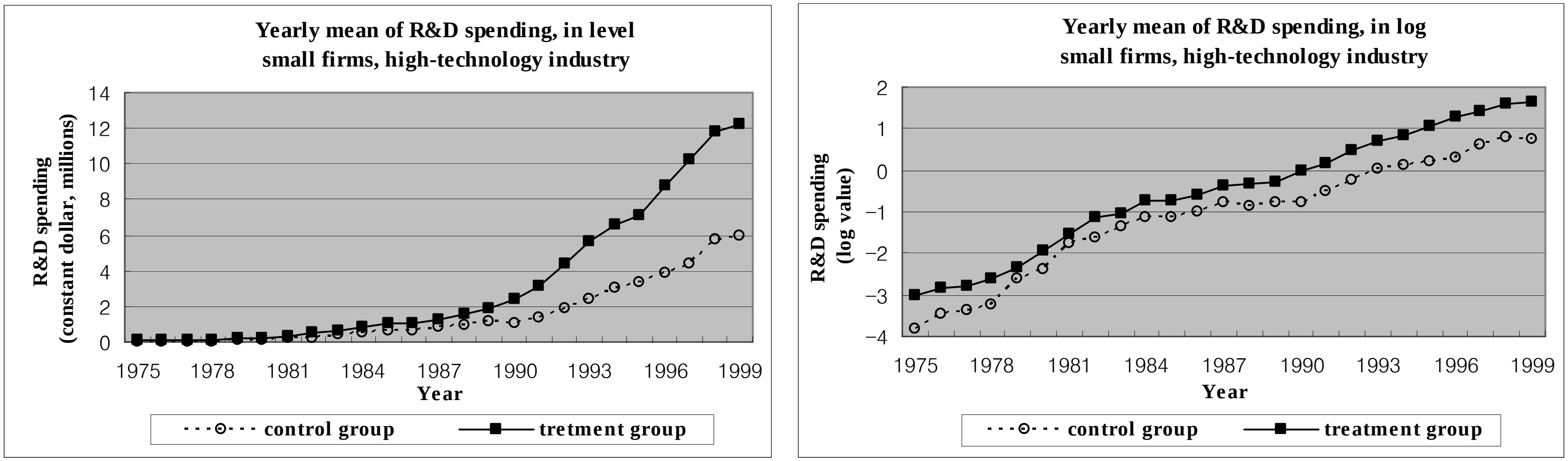


Figure 34. Comparison of historical distribution of R&D spending between the treatment and the control groups for small firms in the high-technology industry

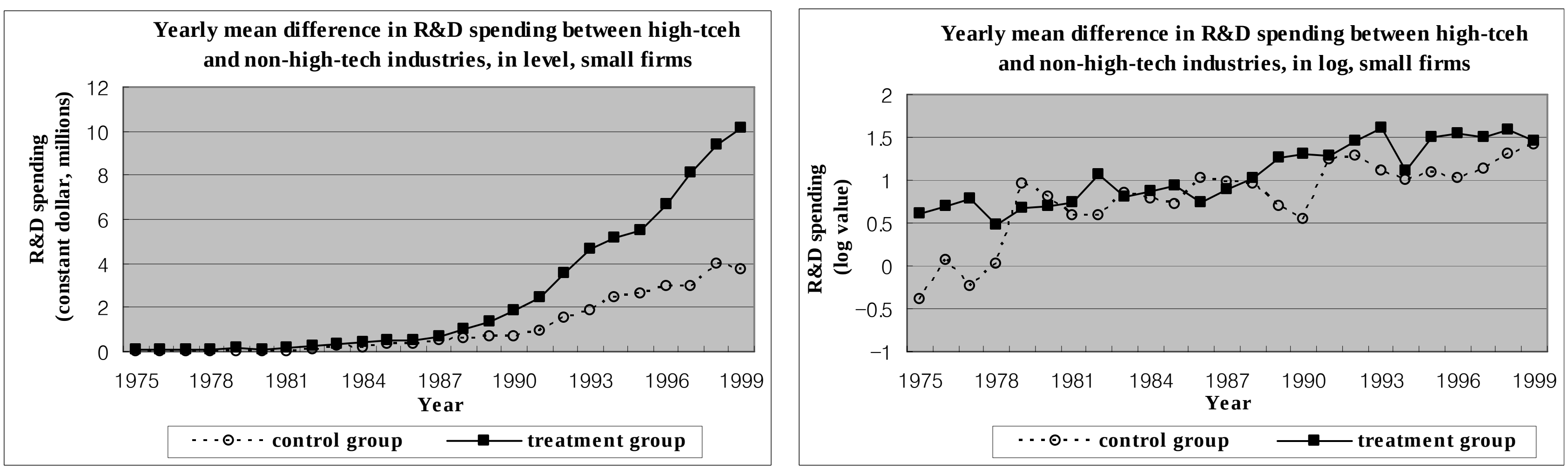


Figure 35. Comparison of historical distribution of R&D spending between the treatment and the control groups for small firms in the high-technology industry (after differencing small firms in the non-high-technology industry)

The differences of R&D spending change over time between the treatment and the control groups, depicted from Figure 30 through 35, can be adjusted by adding some covariates and fixed effects of state, year, and industry in the DD/DDD methods and by finding a matched pair for each observation from selected variables in the matching methods.

Table 50 shows the detailed DD/DDD estimates for different program effects on R&D spending depending on firm size across high-technology industries. The DD estimates are obtained from comparing the treatment group and the control group within the high-technology industry and the DDD estimates are obtained from comparing these two comparison groups within the high-technology industry with adding additional control group, which is the non-high-technology industry. All the estimates in level reveal positive signs and most of them are significant. For small firms, one out of two possible estimates (the DD and DDD estimates) in level is positive and marginally significant. For medium and large firms, both possible estimates in level are positive and significant. The estimates in log show mixed results by each different sized firm. For small firms, the DDD estimate in log is negative and significant. For medium firms, the DD estimate in log is positive and significant. For large firms, the DD/DDD estimates in log are positive and negative respectively. One possible explanation of the smaller outcome increases in the treatment group for small firms relative to the control group, based on the negative and significant estimate by using the DDD method, is that the non-high-technology industry is not a valid comparison group in particular for small sized firms, which is found in Figure 35. In the analysis of R&D spending by firm size across high-technology industries by using DD/DDD methods, most of the statistically significant estimates are obtained from the DD methods.

Table 50. The DD/DDD estimates of effects on R&D spending by firm size across high-technology industries

R&D spending for high-technology industry	Level						Log					
	Small		Medium		Large		Small		Medium		Large	
	DD	DDD	DD	DDD	DD	DDD	DD	DDD	DD	DDD	DD	DDD
T.post	0.12* (0.07)	0.11 (0.08)	1.29** (0.58)	0.77* (0.41)	8.89** (4.35)	7.19* (3.83)	-0.01 (0.03)	-0.20* (0.11)	0.13** (0.05)	0.05 (0.08)	0.009 (0.018)	-0.01 (0.02)
Sales (lagged)	0.006* (0.001)	0.006** (0.001)	0.04** (0.01)	0.03** (0.01)	0.003** (0.001)	0.002** (0.001)	0.02** (0.006)	0.02** (0.006)	-0.01 (0.02)	-0.01 (0.01)	0.02* (0.01)	0.05** (0.01)
Unemployment rate	0.03** (0.01)	0.03** (0.01)	-0.04 (0.09)	-0.03 (0.05)	-0.24 (0.30)	-0.21 (0.15)	0.09** (0.04)	0.08** (0.04)	-0.04 (0.06)	-0.06 (0.04)	0.003 (0.02)	0.002 (0.01)
Labor force	0.00009 (0.0001)	0.00008 (0.0001)	0.0001 (0.0004)	-0.0001 (0.0003)	0.002 (0.004)	0.002 (0.003)	0.03* (0.01)	0.03** (0.01)	-0.02 (0.02)	-0.0001 (0.02)	0.01 (0.01)	0.006 (0.005)
Knowledge Stocks	0.0003 (0.0002)	0.0003 (0.0002)	0.002** (0.001)	0.001** (0.0006)	0.007 (0.005)	0.002 (0.003)	0.01** (0.004)	0.009** (0.002)	0.001 (0.004)	0.002 (0.007)	0.008* (0.004)	0.003 (0.003)
Private R&D spending (lagged)	0.98** (0.03)	0.98** (0.03)	0.63** (0.16)	0.66** (0.16)	0.95** (0.03)	0.96** (0.03)	0.86** (0.01)	0.85** (0.009)	0.87** (0.02)	0.86** (0.01)	0.94** (0.007)	0.93** (0.007)
N	9759	12511	3760	5611	8708	18670	8233	10301	3286	5148	8631	17989
Adjusted R^2	0.7788	0.7809	0.5859	0.5959	0.9685	0.9682	0.8457	0.8403	0.8783	0.8811	0.9647	0.9614

Notes:
· ** Indicates statistical significance at the level of 95%
· * Indicates statistical significance at the level of 90%.
· In the level analysis, private R&D spending is measured in constant 1984 million dollars. The sales are measured in constant 1984 million dollars. Labor force is measured in millions. Knowledge stocks are measured as patent stocks depreciated at a rate of 10%, in units of tens. Unemployment rate is measured in percentage.
· In the log analysis, all variables are transformed by natural log.
· The numbers in parentheses indicate heteroskedasticity-consistent and serial-correlation-consistent standard errors.

I conclude that state R&D tax credits positively affect to increase private R&D spending for all kinds of firms across high-technology industries at least measured by the amount of R&D spending. The findings relatively strongly support the effects for medium firms based on the significant estimates in both level and log, while the findings relatively weakly support the effects for small and large firms based on the marginally significant estimate for small and large firms in level and marginally significant and negative estimate for small firms in log.

The average yearly program effects on amount of R&D spending by firm size are assessed based on either 95% or 90% confidence intervals depending on the existence of significant estimates. First, the average program effects per high-technology firm in a year are assessed in terms of the amount of private R&D spending measured in 1984 constant dollars. For small firms, program effects range from 2,000 to 230,000. For medium firms, program effects range from 80,000 to 2,000,000. For large firms, program effects range from 500,000 to 17,000,000. The effect sizes tend to be larger by the increase in firm size. Then, the average program effects are assessed in terms of the % change of private R&D spending. For small firms, program effects are between -38% and -1%, while for medium firms those are between 4% and 23%.

Within the analysis using DD/DDD methods, I also find some relationships between R&D spending and relevant factors across high-technology industries. Among selected covariates, the lagged sales, knowledge stocks, and the lagged private R&D spending reveal statistically significant and positive relationships with R&D spending for all three kinds of firms. Interestingly, the unemployment rate reveals statistically significant estimates for only small firms and the relationships with R&D spending is positive, implying that small firms might increase their R&D activities even in the worse economic condition. These results in the high-technology industry analysis are mostly similar to the results in the all industry analysis, and therefore it indicates that the overall economic environments affect the growth in private R&D spending in similar ways across high-technology industries and all industries.

The matching estimates of different program effects on R&D spending depending on firm size across high-technology industries are summarized in Table 51. The matching estimates are obtained based on the nearest matching method and two kinds of metrics, which are Mahalanobis metrics and propensity score metrics with caliper. Within propensity score metrics, the matching is conducted either one-to-one matching or multiple matching. Therefore, there are three kinds of matching estimates for each estimate. The significance of matching estimates is estimated by using the bootstrapping method. Most of the matching estimates are positive but, only the estimates in level, for small and large firms, are statistically significant at least at the moderate level. Unlike the DD/DDD estimates, every matching estimate for medium firms is insignificant, showing unclear program effects. This result possibly comes from the lack of comparable observations for medium firms.

The average program effects on R&D spending by firm size across high-technology industries are measured by confidence intervals at 90% level for small firms and at 95% level for large firms depending on the statistically significant estimates. For small firms, the average program effects range from 70,000 to 700,000 1984 constant dollars and for large firms, the effects range from 15,000,000 to 57,000,000 1984 constant dollars.

Table 51. The matching estimates of effects on R&D spending by firm size across high-technology industries

R&D spending for the high-technology industry			# of observations		Estimate	Standard error
			Treatment	Control		
Level, small firms	Mahalanobis metrics		202	101	0.41*	0.20
	Propensity score metrics	Caliper=0.005 # of matched=1	54	101	0.57	0.60
		Caliper=0.005 # of matched=3	54	101	0.78	0.62
Level, medium firms	Mahalanobis metrics		84	30	2.36	1.49
	Propensity score metrics	Caliper=0.005 # of matched=1	5	30	3.85	8.48
		Caliper=0.005 # of matched=3	5	30	3.84	4.80
Level, large firms	Mahalanobis metrics		223	85	36.43**	10.64
	Propensity score metrics	Caliper=0.005 # of matched=1	33	85	15.62	62.98
		Caliper=0.005 # of matched=3	33	85	15.58	45.60
Log, small firms	Mahalanobis metrics		202	101	0.19	0.22
	Propensity score metrics	Caliper=0.005 # of matched=1	29	101	0.02	0.41
		Caliper=0.005 # of matched=3	29	101	0.14	0.60
Log, medium firms	Mahalanobis metrics		84	30	0.27	0.41
	Propensity score metrics	Caliper=0.005 # of matched=1	10	30	0.17	0.65
		Caliper=0.005 # of matched=3	10	30	0.10	0.68
Log, large firms	Mahalanobis metrics		223	85	-0.07	0.17
	Propensity score metrics	Caliper=0.005 # of matched=1	33	85	0.11	0.46
		Caliper=0.005 # of matched=3	33	85	0.11	0.45

Notes
· Matching is performed based on the nearest neighbor matching method.
· Matching variables are lagged R&D spending, lagged sales, knowledge stocks, labor forces, firm size, per capita income, share of high school graduates, share of college graduates, SBIR funds, university funds, population change, crime rate, poverty rate, unemployment rate, and manufacturing establishment density.
· All dollar values are converted in constant 1984 dollars. R&D spending is measured in millions.
· Standard errors are estimated by using the bootstrapping method.

In sum, the effects of state R&D tax credits on R&D spending across high-technology industries are not limited to large firms, instead are widespread to all three sizes of firms. These findings are supported by both analytical methods. In the DD/DDD methods, the estimates for medium firms are the most reliable, while in the matching methods those are not, based on no significant estimate.

In this section, I examined program effects on R&D spending by firm size. The positive effects from state R&D tax credits in terms of increasing firm's R&D spending are detected from all three kinds of firms, at both industrial levels and from both analytical methods. It tells us that state R&D tax credits broadly generate positive increases in private R&D spending, across different industries and across different sized firms. By construction, the state R&D tax credits are eligible for every qualified R&D expenditure and therefore, the broadly spread-out effects from the program can be resulted in. This study provides an evidence for these possible extensive effects. These findings also tell us that even small firms could gain the benefits from tax credits, which is opposite to the common beliefs. The estimated effects include indirect effects from possible spillovers and catch-up efforts and therefore it implies that even though small firms cannot claim the credits directly they also can receive the benefits from other firms indirectly. The revealed effect sizes are larger by the increase in firm size. In the next section, I examine program effects on employment by firm size.

7.4 ANALYSIS OF THE EFFECTS OF EMPLOYMENT ON GROWTH BY FIRM SIZE

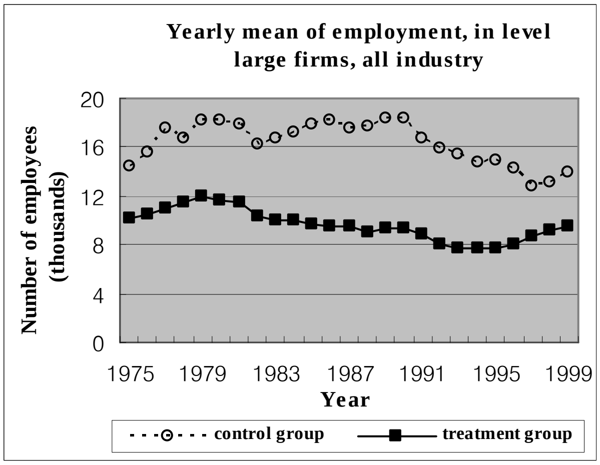
In this section I analyze the effects of state R&D tax credits in terms of different program effects on increasing employment by firm size. The hypothesis to be tested is whether there is any different employment growth resulted from state R&D tax credits depending on firm size. This hypothesis is tested by estimating the average additional increase in employment for states with tax credits before and after the program relative to states with no tax credits for small, medium and large firms separately and then comparing them. Due to data availability, the analysis is performed at the firm level only.

In the above analysis for program effects on increasing overall employment level in section 7.2, the findings are no program effects for the all industry employment and positive program effects for the high-technology industry employment. However, the analysis for program effects by firm size provides the possible distinction of these overall effects across industries. Namely, across all industries, the significant and positive increases in employment for large firms are detected, while no significant increases in employment for small and medium firms are detected, from both the DD and the matching methods. Across high-technology industries, the significant and positive increases in employment for medium and large firms are found, while no significant increases in employment for small firms are found, from either the DD/DDD or the matching methods. These results suggest that positive effects on employment growth for larger firms are dominant for both industrial levels while there is no effect on employment growth for small firms across any industries. In other words, state R&D tax credits generate the increase in employment in particular for larger firms while these policy programs do not generate the increase in employment for small firms. The detailed estimates of the program effects on employment growth by firm size follow in the next section.

7.4.1 Employment by firm size for the all industry

For analyzing program effects on increasing employment by firm size across all industries, the program effects are estimated separately by three-sized firms and then compared each other. Figures 36, 37, and 38 compare the historical distribution of yearly means of employment between two comparison groups for the all industry, by firm size and in level and log values. The historical distribution of yearly means of employment by firm size for the all industry follows mostly similar patterns over time. By dividing comparison groups based on firm size, two comparison groups have more similar characteristics for increasing their employment over time, providing the ground of better comparison, compared to overall assessment by using all observations of the firm-level data (in the section 7.2, Figure 25). Again, because of the existence of substantially large sized firms in the control group, the yearly means of employment for large firms are higher in the control group in both level and log analyses.

The empirical analyses could reduce possible systematic differences between two comparison groups by adding some covariates and fixed effects for year, state, and industry in the DD/DDD methods and by finding the closest matched pairs in the matching methods.



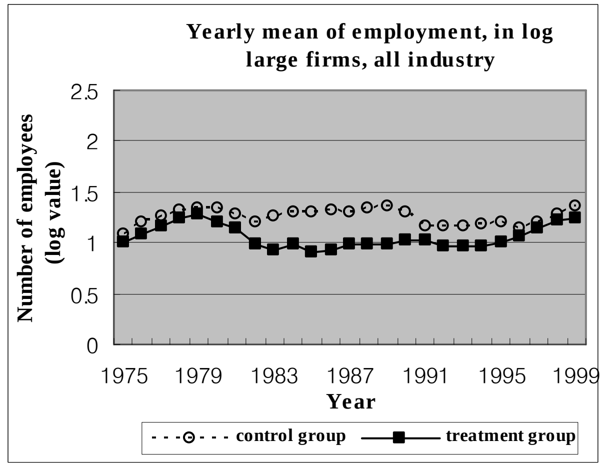
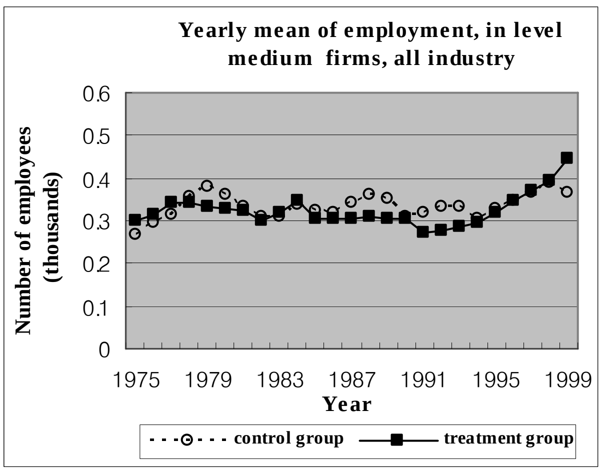


Figure 36. Comparison of historical distribution of employment between the treatment and the control groups for large firms in the all industry



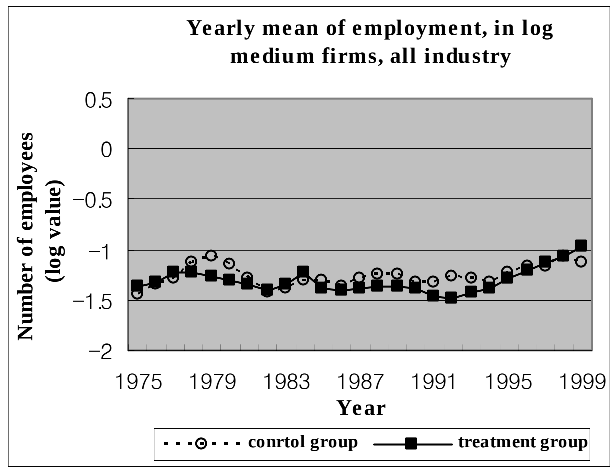
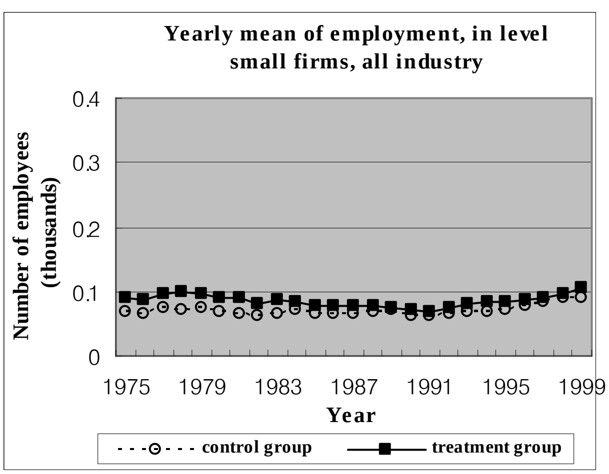


Figure 37. Comparison of historical distribution of employment between the treatment and the control groups for medium firms in the all industry



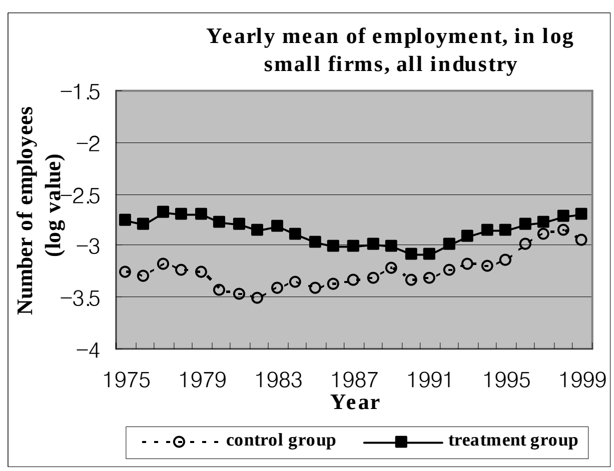


Figure 38. Comparison of historical distribution of employment between the treatment and the control group for small firms in the all industry

In detail, the DD estimates of the effects on employment by firm size across all industries are represented in Table 52. The DD estimate in level is positive and marginally significant for large firms while they are insignificant for small and medium firms. Most of the DD estimates in log are insignificant. These results tell us that large firms in the all industry experienced the increase in their employment from state R&D tax credits, while small and medium firms did not. The average yearly program effects on increasing employment can be measured by the 90% confidence intervals solely for large firms, based on the significant estimate. For large firms, state R&D tax credits have the effect of increasing about 30~900 employees.

Table 52. The DD estimates of effects on employment by firm size for the all industry

Employment for the all industry	Level			Log		
	Small	Medium	Large	Small	Medium	Large
T.post	0.001 (0.002)	0.01 (0.007)	0.48* (0.27)	0.02 (0.03)	0.03 (0.02)	-0.008 (0.011)
Sales (lagged)	0.0005* (0.0003)	0.0002** (0.0001)	-0.00001 (0.0001)	0.02** (0.006)	0.01 (0.01)	0.04** (0.01)
Private R&D spending	0.0005 (0.0003)	0.0004 (0.0003)	-0.001 (0.003)	0.01** (0.005)	0.02** (0.009)	0.02** (0.004)
Labor force	0.0008** (0.0002)	0.004** (0.001)	0.007 (0.03)	0.02** (0.008)	0.01 (0.01)	0.0008 (0.004)
Knowledge Stocks	-0.00001 (0.00002)	-0.0001 (0.0001)	-0.001 (0.001)	-0.001 (0.002)	-0.008* (0.004)	-0.0001 (0.0009)
Employment (lagged)	0.83** (0.02)	0.84** (0.02)	0.99** (0.02)	0.87** (0.01)	0.78** (0.01)	0.89** (0.007)
N	11566	5344	18504	9642	4929	17867
Adjusted R^2	0.7493	0.7261	0.9847	0.8061	0.7409	0.9758

Notes: · ** Indicates statistical significance at the level of 95%
· * Indicates statistical significance at the level of 90%.
· In the level analysis, private R&D spending and federal funding are measured in constant 1984 million dollars. Employment is measured in thousands. The GSP, the PI, and the sales are measured in constant 1984 million dollars. Labor force is measured in millions. Knowledge stocks are measured as patent stocks depreciated at a rate of 10%, in units of tens. Unemployment rate is measured in percentage.
· In the log analysis, all variables are transformed by natural log.
· The numbers in parentheses indicate heteroskedasticity-consistent and serial-correlation-consistent standard errors.

The analyses using the DD method include empirical findings for some relationships between employment and relevant factors across all industries. Most of selected covariates reveal the positive and significant relationships with employment growth. In detail, the lagged sales, the lagged private R&D spending, and the lagged employment which are measured at the firm level and labor forces measured at the regional level, show positive relationships with employment growth.

Next, the detailed matching estimates of the effects on increasing employment by firm size across all industries are shown in Table 53. There are three kinds of matching estimates, which are nearest matching estimates either using Mahalanobis metrics or using propensity score metrics with caliper (specifying number of matches as 1 or 3). The significance of estimates is obtained from the bootstrapping method. The matching estimates by firm size show some mixed results, which means some positive ones and some negative ones, and most of them are insignificant. Only the matching estimate in log for large firms is significant. This significant estimate indicates that the yearly average program effect for large firms, measured by 90% confidence intervals, is between 10 and 25 at the percentage term. The revealed effect size from the matching method is big. One possible explanation for this big effect size is that the estimate is obtained from the matching estimates by using Mahalanobis metrics, which means all observations are used for the matching and therefore, there is a possibility of some matched pairs with quite different characteristics, based on the absence of comparable firms in the control group or from the heterogeneity of matched pairs within datasets. Other matching estimates in log for large firms by using propensity score metrics are insignificant and even negative.

Table 53. The matching estimates of effects on employment by firm size for the all industry

Employment for the all industry			# of observations: Treatment	# of observations: Control	Estimate	Standard error
Level, small firms	Mahalanobis metrics		267	152	0.003	0.005
	Propensity score metrics	Caliper=0.005 # of matched=1	67	152	0.02	0.02
		Caliper=0.005 # of matched=3	67	152	0.02	0.02
Level, Medium firms	Mahalanobis metrics		113	58	-0.003	0.04
	Propensity score metrics	Caliper=0.005 # of matched=1	21	58	0.05	0.13
		Caliper=0.005 # of matched=3	21	58	0.05	0.16
Level, large firms	Mahalanobis metrics		398	209	1.39	0.91
	Propensity score metrics	Caliper=0.005 # of matched=1	176	209	-2.11	2.63
		Caliper=0.005 # of matched=3	176	209	-1.28	2.50
Log, small firms	Mahalanobis metrics		267	152	0.01	0.12
	Propensity score metrics	Caliper=0.005 # of matched=1	49	152	0.45	0.39
		Caliper=0.005 # of matched=3	49	152	0.42	0.32
Log, Medium firms	Mahalanobis metrics		113	58	0.02	0.15
	Propensity score metrics	Caliper=0.005 # of matched=1	21	58	-0.003	0.36
		Caliper=0.005 # of matched=3	21	58	-0.01	0.44
Log, large firms	Mahalanobis metrics		398	209	0.26**	0.10
	Propensity score metrics	Caliper=0.005 # of matched=1	131	209	-0.06	0.20
		Caliper=0.005 # of matched=3	131	209	-0.09	0.15

Notes
· Matching is performed based on the nearest neighbor matching method.
· Matching variables are lagged R&D spending, lagged sales, knowledge stocks, labor forces, firm size, per capita income, share of high school graduates, share of college graduates, SBIR funds, university funds, population change, crime rate, poverty rate, unemployment rate, and manufacturing establishment density.
· All dollar values are converted in constant 1984 dollars.
· Standard errors are estimated by using the bootstrapping method.

Consequently, state R&D tax credits generate positive effects on increasing employment for large firms across all industries, while these programs do not for medium and small firms. In the next section, I examine program effects on increasing employment by firm size, across high-technology industries.

7.4.2 Employment growth by firm size for the high-technology industry

For analyzing the effects on increasing employment by firm size across high-technology industries, the yearly changes in employment between the treatment and control groups are compared by firm size. Figures 39 through 44 show the historical changes in yearly mean employment of the treatment and the control groups across high-technology industries in level and log, for small, medium and large firms respectively. Unlike above analyses of effects on employment by firm size across all industries, the historical distributions of yearly means of employment by firm size across high-technology industries, which are Figures 39, 41, and 43, do not follow similar patterns especially in level values. It means that the comparison of employment for the high-technology industry cannot be improved by separating comparison groups by firm size. Rather, adding a control group which is non-high-technology industry, which are shown in Figures 40, 42, and 44, results in the better comparison especially for medium and small firms in level values. It tells us that the non-high-technology industry group can be a valid comparison group for capturing the other relevant factors in the analysis of employment growth by firm size across high-technology industries. Again, because of the existence of substantially large sized firms in the control group, the yearly means of R&D spending for large firms are higher in the control group in both level and log analyses.

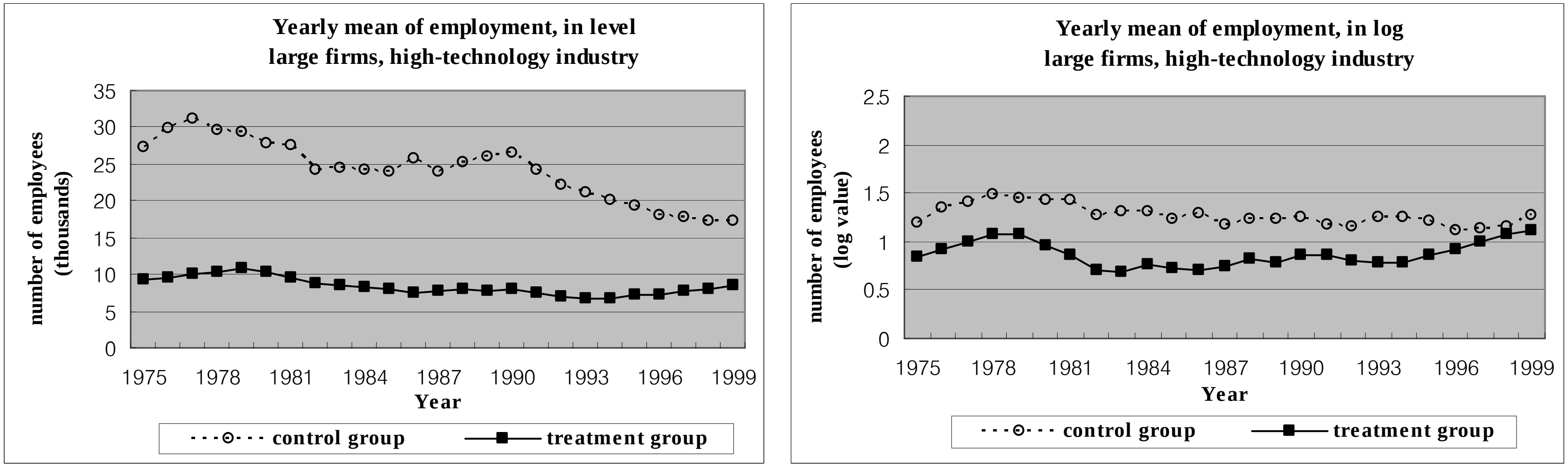


Figure 39. Comparison of historical distribution of employment between the treatment and the control groups for large firms in the high-technology industry

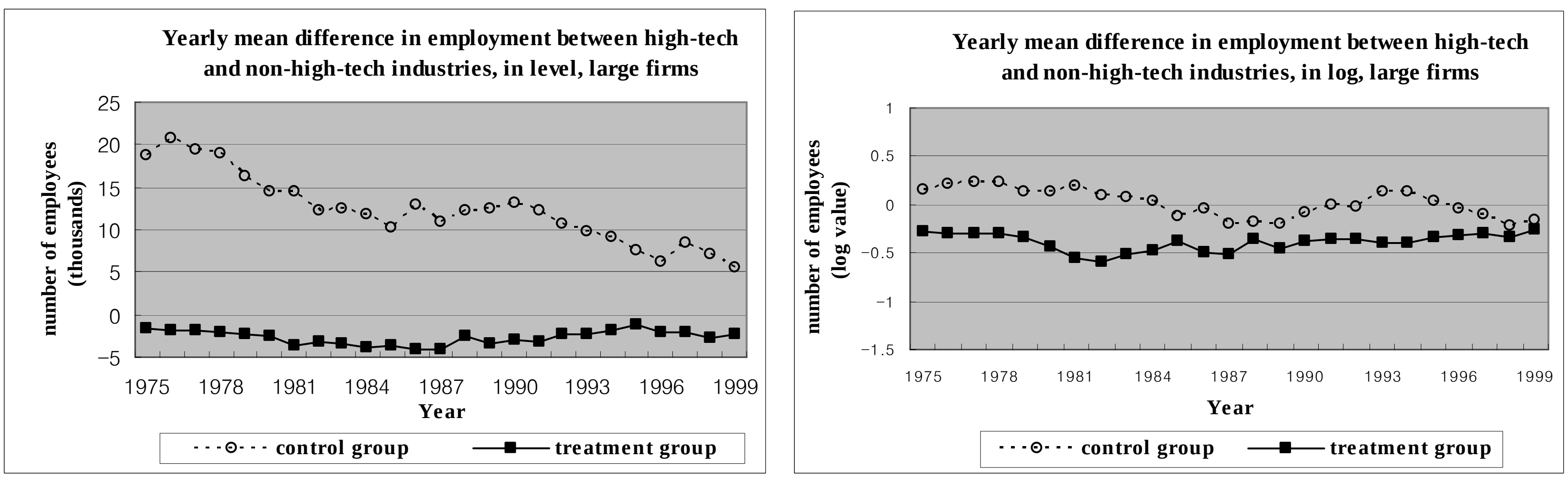


Figure 40. Comparison of historical distribution of employment between the treatment and the control groups for large firms in the high-technology industry (after differencing large firms in non-high-technology industry)

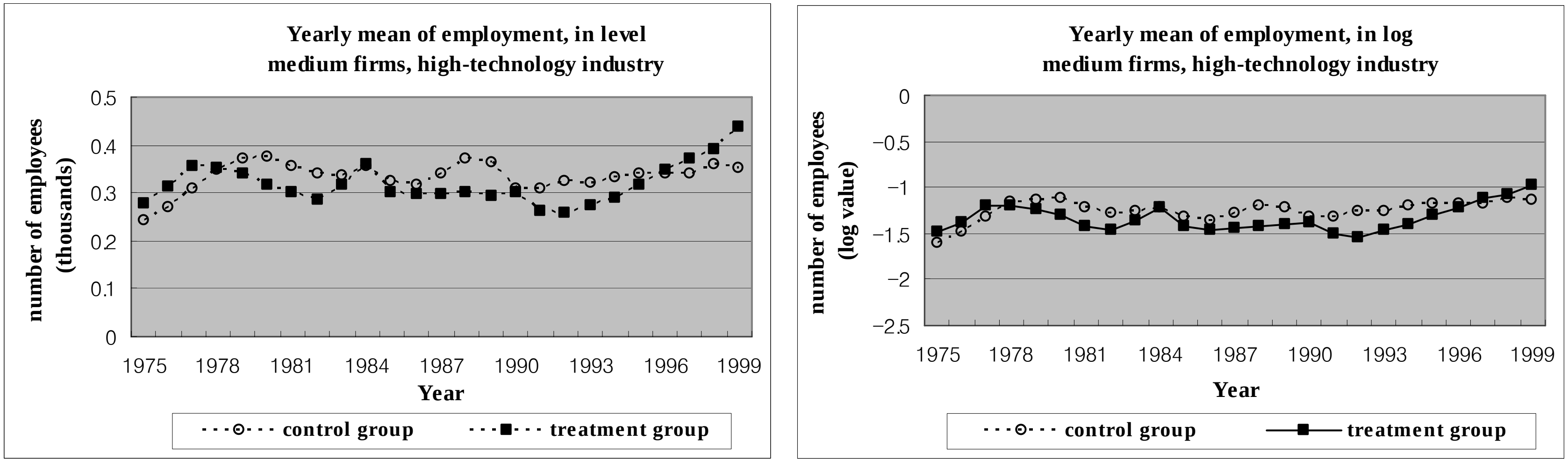


Figure 41. Comparison of historical distribution of employment between the treatment and the control groups for medium firms in the high-technology industry

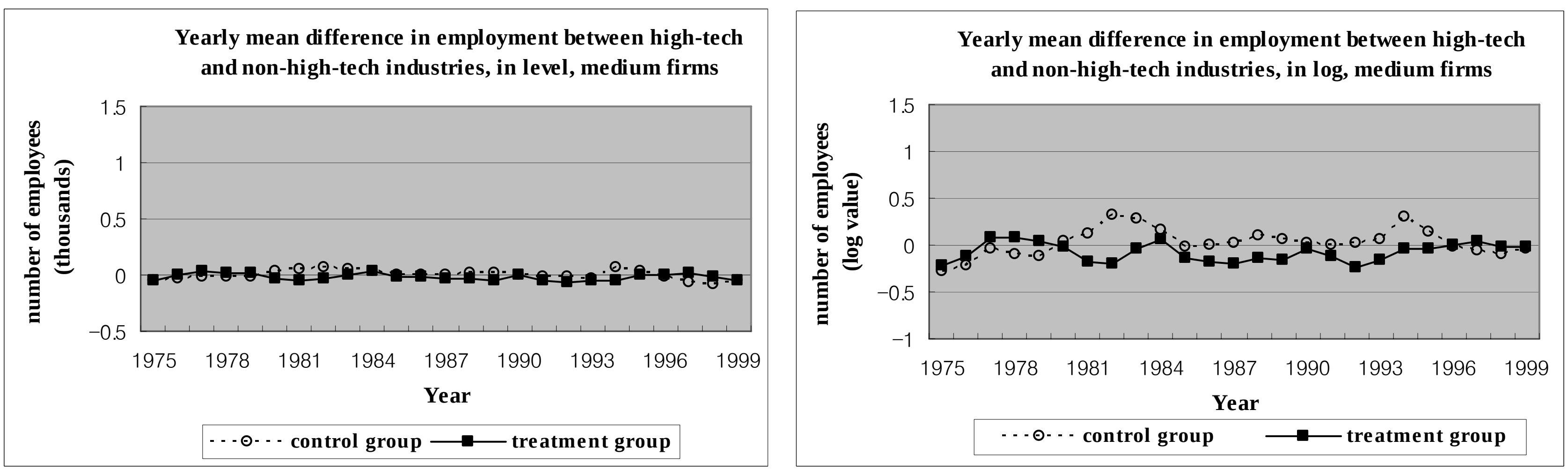
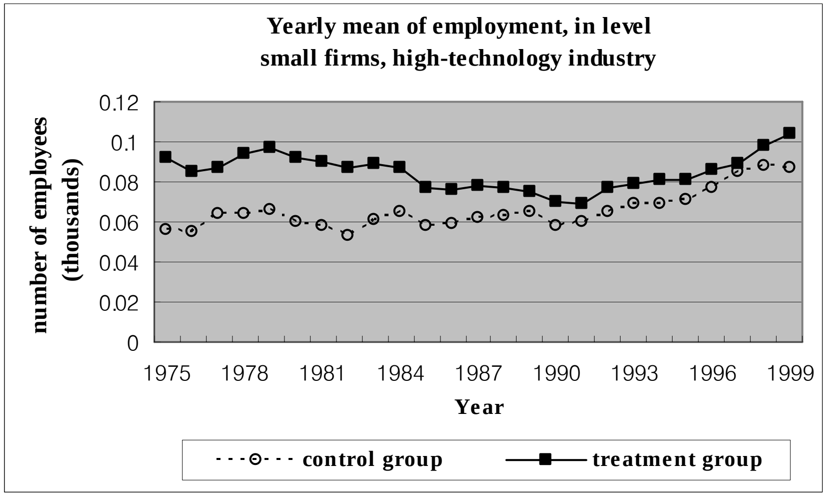


Figure 42. Comparison of historical distribution of employment between the treatment and the control groups for medium firms in the high-technology industry (after differencing medium firms in non-high-technology industry)



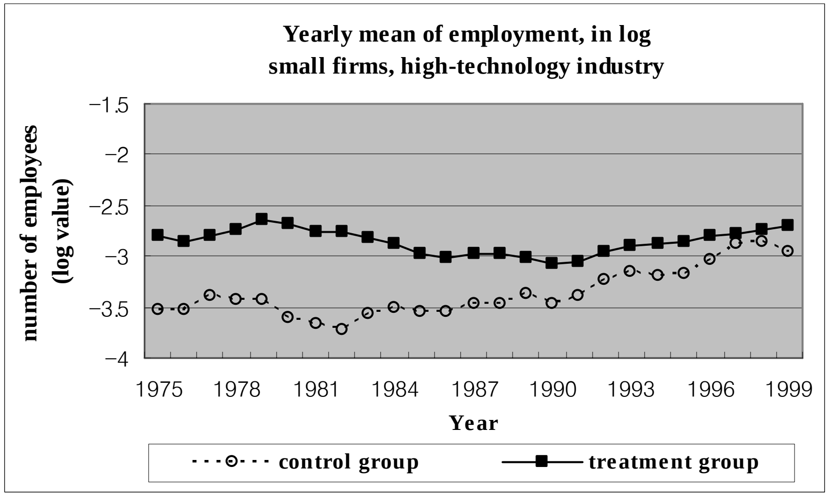
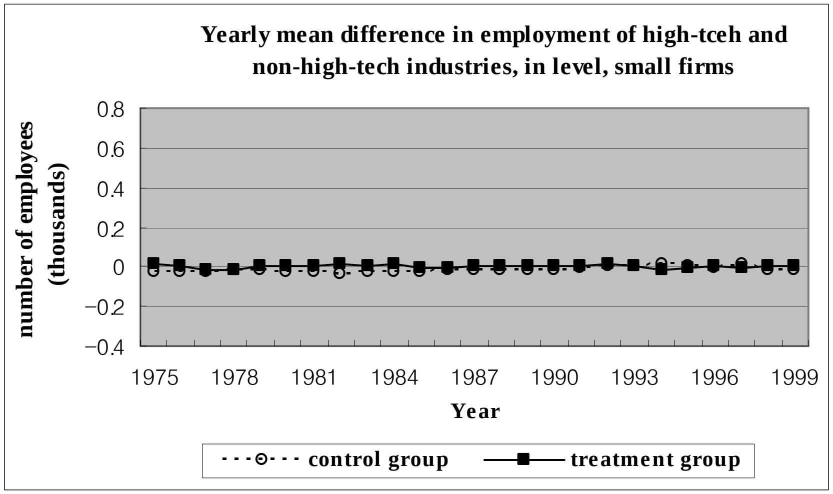


Figure 43. Comparison of historical distribution of employment between the treatment and the control groups for small firms in the high-technology industry



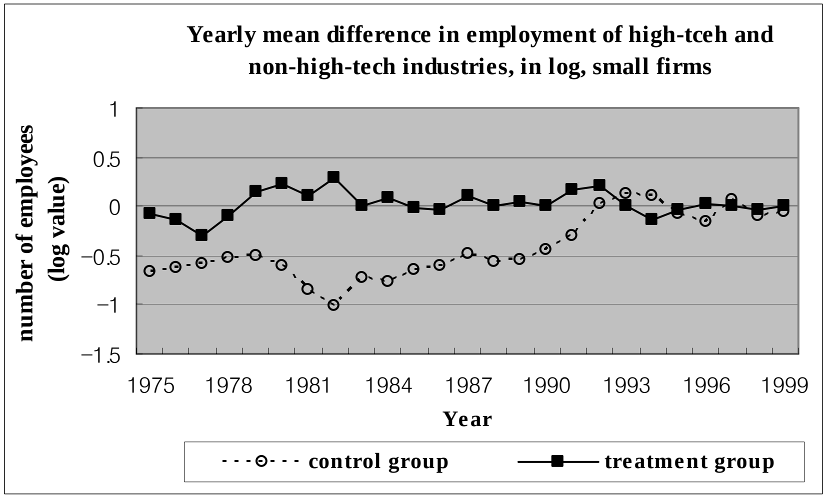


Figure 44. Comparison of historical distribution of employment between the treatment and the control groups for small firms in the high-technology industry (after differencing small firms in non-high-technology industry)

The differences of employment growth over time between the treatment group and the control group, detected from Figures 39 through 44 can be adjusted by adding some covariates and fixed effects of state, year, and industry in the DD/DDD specification and by finding a matched pair for each observation from observable variables.

The detailed DD/DDD estimates for program effects on employment growth by firm size across high-technology industries are summarized in Table 54. There are significant and positive estimates for medium and large firms while no significant estimates for small firms. These results tell us that medium and large firms in the high-technology industry experienced the increase in their employment from state R&D tax credits, while small firms did not. As same as the all industry level analysis, it is revealed that small firms do not receive any benefit from credits for increasing their employment. Based on the statistically significant estimates, the average effects of R&D tax credits per high-technology firm in a year range from 40 to 900 employees for large firms, from 3 to 30 employees for medium firms.

The analyses by using the DD/DDD method include empirical findings for some relationships between employment and relevant factors across high-technology industries. Most of selected covariates, which are the lagged sales, the lagged private R&D spending, and the lagged employment level at the firm level and labor forces at the regional level, reveal positive relationships with employment growth. It tells us that the employment of the high-technology firm is positively affected by the overall level of investment capacity, R&D spending and the previous employment for a firm and the overall regional level of human capital.

Table 54. The DD/DDD estimates of effects on employment by firm sizes across high-technology industries

Employment for high-technology industry	Level						Log					
	Small		Medium		Large		Small		Medium		Large	
	DD	DDD	DD	DDD	DD	DDD	DD	DDD	DD	DDD	DD	DDD
T.post	0.0008 (0.003)	-0.001 (0.002)	0.02** (0.008)	0.02** (0.01)	0.47** (0.24)	0.30 (0.32)	0.02 (0.02)	-0.02 (0.04)	0.04 (0.02)	-0.005 (0.04)	-0.004 (0.01)	0.01 (0.02)
Sales (lagged)	0.004** (0.0008)	0.005* (0.003)	0.004** (0.001)	0.003** (0.001)	-0.0001 (0.0001)	-0.0009 (0.01)	0.02** (0.01)	0.02** (0.008)	0.008 (0.011)	0.01 (0.01)	0.03** (0.01)	0.04** (0.01)
Private R&D spending (lagged)	0.0005* (0.0003)	0.0005 (0.0003)	0.0003 (0.0003)	0.0004 (0.0003)	-0.004 (0.003)	-0.001 (0.003)	0.01* (0.007)	0.01** (0.005)	0.02* (0.01)	0.02** (0.009)	0.02** (0.006)	0.02** (0.004)
Labor force	0.0006* (0.0003)	0.0008** (0.0002)	0.005** (0.001)	0.004** (0.001)	-0.006 (0.03)	0.009 (0.03)	0.02** (0.01)	0.02** (0.006)	0.02* (0.01)	0.01 (0.01)	0.001 (0.008)	0.001 (0.004)
Knowledge Stocks	-0.0001 (0.0001)	-0.0001 (0.0002)	-0.003** (0.001)	-0.001 (0.001)	0.02 (0.03)	-0.01 (0.01)	-0.001 (0.001)	-0.001 (0.002)	-0.008** (0.004)	-0.008* (0.004)	0.001 (0.001)	-0.0001 (0.0009)
Employment (lagged)	0.83** (0.03)	0.83** (0.02)	0.82** (0.02)	0.84** (0.02)	1.02** (0.01)	0.99** (0.02)	0.88** (0.02)	0.87** (0.01)	0.78** (0.02)	0.78** (0.02)	0.90** (0.01)	0.89** (0.007)
N	8951	11566	3523	5344	8518	18504	7616	9642	3381	4929	8447	17867
Adjusted R^2	0.7500	0.7492	0.7280	0.7266	0.9910	0.9847	0.8035	0.8061	0.7370	0.7407	0.9747	0.9758

Notes:
· ** Indicates statistical significance at the level of 95%
· * Indicates statistical significance at the level of 90%.
· In the level analysis, private R&D spending and federal funding are measured in constant 1984 million dollars. Employment is measured in thousands. The GSP, the PI, and the sales are measured in constant 1984 million dollars. Labor force is measured in millions. Knowledge stocks are measured as patent stocks depreciated at a rate of 10%, in units of tens. Unemployment rate is measured in percentage.
· In the log analysis, all variables are transformed by natural log.
· The numbers in parentheses indicate heteroskedasticity-consistent and serial-correlation-consistent standard errors.

Next, the detailed matching estimates of program effects on increasing employment by firm size across high-technology industries are shown in Table 55. There are three kinds of matching estimates, which are nearest matching estimates either using Mahalanobis metrics or using propensity score metrics with caliper (specifying number of matches as 1 or 3). The significance of estimates is obtained from the bootstrapping method. The matching estimates are mostly positive, but only estimates for large firms are marginally significant. Based on the statistically significant estimates, the average effects of R&D tax credits per high-technology firm in a year, measured by 90% confidence intervals, range from 300 to 3,700 employees or between 9 and 56 at the percentage term for large firms. The revealed effect size from the matching method are quite big and wide as like the above analysis of effects on employment growth by firm size across all industries in section 7.4.1. One possible explanation for this big effect size is that the estimate is obtained from the matching estimates by using Mahalanobis metrics, which means all observations are used for the matching and therefore, there is a possibility of some matched pairs with quite different characteristics, based on the absence of comparable firms in the control group or from the heterogeneity of matched pairs within datasets. Other matching estimates in log for large firms by using propensity score metrics are insignificant and even negative.

In sum, the state R&D tax credits affect to increase employment level for medium and large firms across high-technology industries while these programs do not affect to increase employment level for small firms.

Table 55. The matching estimates of effects on employment by firm sizes across high-technology industries

Employment for the high-technology industry			# of observations		Estimate	Standard error
			Treatment	Control		
Level, small firms	Mahalanobis metrics		207	116	0.002	0.006
	Propensity score metrics	Caliper=0.005 # of matched=1	51	101	0.003	0.023
		Caliper=0.005 # of matched=3	51	101	0.008	0.022
Level, medium firms	Mahalanobis metrics		84	31	-0.05	0.06
	Propensity score metrics	Caliper=0.005 # of matched=1	5	31	0.14	0.16
		Caliper=0.005 # of matched=3	5	31	0.10	0.18
Level, large firms	Mahalanobis metrics		225	86	2.04*	1.02
	Propensity score metrics	Caliper=0.005 # of matched=1	65	86	1.93	2.88
		Caliper=0.005 # of matched=3	65	86	1.85	2.67
Log, small firms	Mahalanobis metrics		207	116	0.03	0.12
	Propensity score metrics	Caliper=0.005 # of matched=1	34	116	0.25	0.36
		Caliper=0.005 # of matched=3	34	116	0.25	0.30
Log, medium firms	Mahalanobis metrics		84	31	-0.19	0.20
	Propensity score metrics	Caliper=0.005 # of matched=1	3	31	0.38	0.60
		Caliper=0.005 # of matched=3	3	31	0.26	0.69
Log, large firms	Mahalanobis metrics		225	86	0.33*	0.14
	Propensity score metrics	Caliper=0.005 # of matched=1	33	86	0.04	0.39
		Caliper=0.005 # of matched=3	33	86	0.05	0.36

Notes
· Matching is performed based on the nearest neighbor matching method.
· Matching variables are lagged R&D spending, lagged sales, knowledge stocks, labor forces, firm size, per capita income, share of high school graduates, share of college graduates, SBIR funds, university funds, population change, crime rate, poverty rate, unemployment rate, and manufacturing establishment density.
· All dollar values are converted in constant 1984 dollars.
· Standard errors are estimated by using the bootstrapping method.

From the analysis of the program effects on employment growth by firm size across high-technology industries discussed in this section, the effects on employment growth turn out to be limited to the larger firms for both industrial levels (i.e. the all industry and the high-technology industry). These findings suggest that relatively small amount of R&D spending increases for small firms might make the weak linkage with employment increases, while the relatively sufficient increases of R&D spending for larger firms could make the better linkage with their employment increases. At the other hand, the larger firms are much flexible to adjust their employees while the smaller firms are not. In the next section, the empirical findings in this study are summarized.

7.5 SUMMARY

In this section, I summarize the above analyses of the effectiveness of state R&D tax credit programs on increasing R&D spending and employment, depending on industrial levels, observational levels, and analytical methods. Industrial levels are defined as the all industry and the high-technology industry. The all industry is constructed as manufacturing and service industries, and the high-technology industry is defined as the industry having higher portion of R&D-related employments at the 3-digit SIC level, listed in Chapter 6, Table 20. Observational levels are defined as the state level and the firm level. The selected analytical methods are the DD/DDD methods and the matching methods which are suggested as quasi-experimental ways. The DD method is applied to all of the analyses within this study and the DDD method is applied to the analysis for the high-technology industry. The matching method is applied to the firm level analysis.

Table 56 indicates which outcome changes with state R&D tax credits are larger than the corresponding changes without credits depending on the above analytical categorizations. Within each category, there are several estimates by using alternative methods, alternative covariates, and alternative matching criteria. Therefore, overall assessment is made by the signs of coefficients and the existences of significant coefficients.

Based on the summary of analysis as shown in Table 56, I examine the effects on outcome increases at industrial levels. The most reliable and strongest evidence is the increase in R&D spending across all industries. Namely, I found the larger increases in private R&D spending for states with credits before and after the program relative to states with no credits, from all possible estimates by using both observational levels and both analytical methods. Then, I also found the larger increase in R&D spending across high-technology industries and the larger increase in employment across high-technology industries for states with credits, only from the firm level analysis, indicating that these are relatively weak evidences.[120] However, I found no program effect on the increases in employment across all industries from all possible estimates. Thus, I conclude that the R&D tax credits generate the positive effects on R&D spending which are widespread across all industries, and the positive effects on employment which are limited to the high-technology industry.

This evidence strongly supports the appropriateness of this credit in the ground that the primary purpose of this credit law, which is increasing private R&D spending, has been fulfilled. The limited effects on employment are quite reasonable, in the sense that R&D tax credits aims to support private R&D activities and therefore the primary beneficiary for increasing employment would be the high-technology industry, which has a higher portion of R&D-relevant employees.

---

[120] In case of R&D spending for the high-technology industry, there are only two possible estimates instead of three possible estimates because we don't have the state level data.

Table 56. The summary of analysis

			DD/DDD	Matching
R&D spending	All industry	State level	***Positive, significant***	n/a
		Firm level	***Positive, significant***	***Positive, significant***
	High-technology industry	State level	n/a	n/a
		Firm level	***Positive, significant***	***Positive, significant***

			DD/DDD	Matching
Employment	All industry	State level	Positive, insignificant	n/a
		Firm level	Positive, insignificant	Mixed, insignificant
	High-technology industry	State level	Mixed, insignificant	n/a
		Firm level	***Positive, significant***	***Positive, significant***

Then, the effects on outcome increases by firm size are summarized in Table 57, in which overall signs of estimates and their significances are indicated by firm size. Firm size is defined as small, medium and large, based on the number of employees. The analysis is conducted by using the firm-level data and using both analytical methods.

First, I found different program effects on outcome increases by firm size. For increasing R&D spending, state R&D tax credits generate broad effects for small, medium and large firms, while for increasing employment these programs generate limited effects for larger firms. These findings are clearer in the analysis for the high-technology industry, however, the overall assessments are consistently made at two industrial levels.

Second, the revealed program effects on R&D spending in disaggregated analyses by firm size are also consistent with those in overall industrial level analyses. Namely, I found significant and positive effects on R&D spending across all different sized firms in all and high-technology industries, and I also found these positive effects in both overall industrial level analyses. However, in case of employment, the revealed program effects in disaggregated analyses by firm size are not consistent with those in overall industrial level analyses. Namely, I found significant and positive effects on employment for larger firms in all and high-technology industries, but I found these positive effects only in overall industrial level analysis for the high-technology industry as shown in Table 56. These results tell us that the program effects on employment are relatively more limited in the all industry (only large firms) while these effects are relatively less limited in the high-technology industry (including large and medium firms).

In sum, the above findings indicate that larger firms have received the benefits for increasing both R&D spending and employment, while small firms have received the benefits for increasing R&D spending only. The weak evidence of increasing employment for small firms could be explained as that small firms are less flexible for adjusting employment than large firms and also the amount of R&D spending increases through the credits are relatively small. Based on different investment capacities, adjusting manpower is much harder for small firms while adjusting R&D spending is relatively easy for both small firms and large firms. It is also true that the tax credits are given based on the amount of firm's R&D spending and the R&D spending growth partially results in the subsequent employment growth. Therefore the larger firms could receive the larger amount of benefits and make the better linkage with the increase of employment. Importantly, on the contrary to common beliefs, state R&D tax credits make the positive effects for increasing R&D spending for small firms.

Table 57. The summary of analysis by firm size and outcome

			DD/DDD	Matching
R&D spending	All industry	Small	***Positive, significant***	***Positive, significant***
		Medium	***Positive, significant***	***Positive, significant***
		Large	***Positive, significant***	***Positive, significant***
	High-technology industry	Small	***Positive, significant***	***Positive, significant***
		Medium	***Positive, significant***	Positive, insignificant
		Large	***Positive, significant***	***Positive, significant***
Employment	All industry	Small	Positive, insignificant	Positive, insignificant
		Medium	Positive, insignificant	Mixed, significant
		Large	***Positive, significant***	Mixed, significant
	High-technology industry	Small	Mixed, insignificant	Positive, insignificant
		Medium	***Positive, significant***	Positive, insignificant
		Large	***Positive, significant***	***Positive, significant***

8.0 SUMMARY AND POLICY IMPLICATIONS

This last chapter includes the conclusion and summary of this study, policy implications, major contributions and limitations of this study, and future suggestions. Policy implications are suggested based on major findings and methodologies applied in this study. Then, major contributions and limitations of this study are described and concluding remarks follow.

The primary policy implication of this study is that the utilization of state R&D tax credits is recommended as a generator of R&D spending and employment growth. This research found that state R&D tax credits are meaningful for encouraging private R&D for economic growth for the all industry including the high-technology industry. In addition, this research found that state R&D tax credits are also meaningful for accomplishing regional-level economic growth in terms of the increase in R&D spending and employment. The effects on R&D spending are spread across the all industry, while the effects on employment are limited to the high-technology industry.

As a tax incentive for promoting private R&D, the state R&D tax credit programs can increase productivity and encourage innovation, then accomplish economic growth in the long run. Supporting private R&D is expected to produce valuable benefits for society as a whole on the grounds that R&D can produce higher social returns to scale than the private returns that occur through spillovers. In particular, the broad effects of state R&D tax credits on R&D spending for the all industry could generate a broad externality effect for society in general.

Because these positive externalities are believed to be localized with some geographical barriers, supporting private R&D at the regional level is especially meaningful. As an indirect government intervention, state R&D tax credits have some advantages, such as providing flexibility and autonomy to private firms without distorting the firms' decision making, subsequently generating efficient resource allocation. Therefore, state R&D tax credits can be utilized along with other complementary direct intervention policies to encourage private R&D.

The other important policy implication of this study is that the effects of state R&D tax credits vary depending on outcome and firm size. State R&D tax credits had positive effects on R&D spending growth for firms of all sizes including small firms. Contrary to the common belief that large firms experience limited benefits from state R&D tax credit programs, this study provided strong evidence of broad program effects to all kinds of firms across industries. This finding has significant policy implications in the sense that supporting R&D activities for small firms is an important objective of R&D-relevant policies due to the large fixed cost generally involved in R&D activities. Meanwhile, this research found that the programs only had positive effects on employment growth for larger firms. Therefore, the utilization of supplementary policies is recommended for increasing employment, particularly for small firms.

This research contributed to the evaluation of public policy by developing evaluation strategies based on a quasi-experimental approach. The selected program evaluated in this study was the state R&D tax credit, which is one of regional economic development policies as well as R&D-relevant policies. With the current lack of quantitative evaluation of regional development policies and R&D-relevant policies, this study can be seen as an important exemplar in this field. The empirical findings and corresponding policy implications provide valuable knowledge in the academic field as well as in the practical field. More importantly, the evaluation methods and the detailed evaluation strategies developed in this study can be applied for evaluating other regional level policies or other R&D-relevant policies.

The limitations of this study relate to the methods and datasets. First, the evaluation methods applied in this study evaluated the overall average effects of state R&D tax credits across states and industries. Therefore, there was no analysis of the sensitivity to the credit rate or of program effect by state or industry. Second, extended datasets are required for separating direct effects from indirect effects and for providing additional empirical evidence for the effects of R&D tax credit programs on R&D spending for the high-technology industry.

This study attempted to develop an evaluation strategy using a quasi-experimental approach. The reliability and credibility of this type of research design highly depends on the availability of current knowledge and the feasibility of available research methods and datasets. Based on the falsification logic originally developed by Popper (1959), this study attempted to find empirical evidence by ruling out rival hypotheses as much as possible.

8.1 METHODOLOGY AND FINDINGS

8.1.1 Major methodology

The purpose of this study was to evaluate state R&D tax credits, as a social program and more specifically as a regional economic development program. I used the quasi-experimental approach to do this because it was difficult to perform a randomized social experiment. Therefore, the most fundamental task was to deal with the naturally imbedded dissimilarity between an experimental group and a control group in a social setting. I accomplished this by constructing rival hypotheses and ruling them out as much as possible. Rival hypotheses examined in this study included the state tax environment, the business environment, the R&D environment, the policy environment, and the firm environment. The specific features within each environment are listed in Chapter 6, Table 27, and their relationships are represented in Chapter 6, Figure 16.

I used two major analytical tools that have been widely applied in recent program evaluations— the difference-in-differences method (DD method) and the matching method. The DD method and the matching method estimate the differences in outcome changes between an experimental group and a control group, which indicates a program effect. The DD method estimates the program effect by comparing the average outcome change of each group assuming the average outcome change of a control group indicates possible change with no program implementation. The matching method estimates the program effect by finding matched observations in the experimental and control groups based on observable characteristics and averaging the outcome differences of each matched pair.

I developed several evaluation strategies for controlling the dissimilarity of the two comparison groups, which is a critical step in the quasi-experimental approach. First, comparison groups were created based on states with tax credits and states without tax credits. Within these groups, states with similar characteristics in the utilization of state R&D tax credits and overall state tax structure were chosen for closer comparison. In addition, I included some covariates in the empirical DD specification, and chose matching variables based on rival hypotheses for capturing systematic differences between comparison groups. Other important strategies for reducing the dissimilarity of comparison groups included using the difference-in-difference-in-differences method (DDD method) and the matching method incorporated with the

DD method. I also used multiple observations with varying time periods for each comparison group, thus controlling for time-specific effects and state/industry-specific effects as well. This type of quasi-experimental research design is called “the interrupted time series design adding the non-equivalent control group and adding switching replications.” Switching replications mean that each state used its own specific tax credit program and that the programs were implemented at different times. This research design is superior to the alternatives, which are “the simple non-equivalent control group design,” which consists of comparing two groups with two observations (one for pretest and one for posttest), and “the simple interrupted time series design,” which consists of using multiple observations varied with time periods but based on the dependent variable of only one experimental group. These strategies were all developed with a linkage of rival hypotheses. In addition, using multiple levels of observations with multiple analytical methods provided more reliable and stronger findings.

However, some rival hypotheses could be remaining in a quasi-experimental setting. I examined these possibility by assessing how evaluation strategies developed in this study deal with rival hypotheses. Some factors in rival hypotheses were still remaining due to the lack of measurable variables. In particular, in the DD/DDD methods, more factors in rival hypotheses were possibly remaining than in the matching methods due to the more limited availability of data for covariates than matching variables.

In detail, three factors, which are possible different level of agglomeration economies, the level of competition and available human capital (highlighted in Chapter 6, Table 39), were recognized as the fundamental factors for deciding the level of private R&D activities especially at the regional level but not being able to be eliminated directly within the model. These factors are relevant to selection (i.e. systematic differences in characteristics between experimental observations and others) among the possible threats of validity. Because private R&D activities are quite skewed across states and across regions, the outcome changes from these possible different environments might combine the outcome change from program and cause to evoke the confusion. Besides the above three factors, there were some factors possibly causing another possible threat to validity, which is history (i.e. possible third factors include other events occurring concurrently). These are mostly under different government policy environment and more specifically other local economic development strategies and other R&D-supporting program either at the federal level or at the state level. It implies that the empirical analysis in this study could not properly capture the possible variations from regional differences in the private R&D performances, which consists of one of significant factors determining the outcome changes.

By re-evaluating rival hypotheses in terms of the possibility of ruling them out within empirical specifications, it is revealed that the fundamental difficulty involved in evaluating state R&D tax credits is the skewness of private R&D performance in the US. In detail, six states have performed approximately 50% of whole R&D activity in US and ten states have done nearly 70% of this activity during the 1990s. This skewness is primarily considered in constructing valid control group by limiting the observations by year and including some states performing a high level of private R&D within control group. However, there is still a room for uncontrolling the possible variations from different overall business and R&D-specific environments.

Within this context, it is also worth mentioning that most of states having a high level of private R&D performance (nine out of ten states) already had state R&D tax credits. It means that instead of whether this program should be adopted or not, how to utilize this program could be more important and practical research question. Therefore, with an empirical finding of overall positive effects of the programs across states, it is necessary to provide the detailed analysis for the better implementation, such as different effects by firm size or different effects by the different utilization rules.

8.1.2 Major findings

Using the research design described above, this study analyzed the effectiveness of state R&D tax credits. As a prerequisite, the detailed features of various state R&D tax credits and relevant state corporate tax systems were examined in Chapter 4, section 4.3. Based on the analyses and evaluation strategies mentioned above, the effects of state R&D tax credits on R&D spending and employment were estimated at the state level and the firm level, for the all industry and the high-technology industry respectively. In addition, the effects were estimated separately by firm size. These estimates provided valuable knowledge for establishing regional level R&D policy.

The major findings of this study are 1) state R&D tax credits increased R&D spending and employment, 2) positive effects on R&D spending were widespread across industry while positive effects on employment were limited to the high- technology industry, 3) positive effects on R&D spending were widespread across firms of all sizes in both all and high-technology industries, and 4) positive effects on employment were mostly absorbed into large firms in both all and high-technology industries. Importantly, the effects estimated in this study included direct effects to the firms that received the credits, and indirect effects to firms that did not receive the credits but might benefit from sub-contracts, spillover effects, or catching-up efforts.

Based on the above findings, I conclude that state R&D tax credits have been successful. This study provides empirical evidence of the larger increase in R&D spending and employment in states with credits compared with states without credits. Specifically, the effects on R&D spending tended to be widespread across industries while the effects on employment tended to be limited to the high-technology industry. These findings became more reliable and stronger based on multiple methods and multiple datasets. The effects also became more precise to control for possible relevant effects from observable and unobservable variables based on rival hypotheses.

Along with the above overall assessment, I also found some interesting distinctions of program effects by industry and firm size. First, the increase in R&D spending was extensively detected in the all industry (including the high-technology industry) while the increase in employment was found on a limited basis within the high-technology industry. Second, in the same way, the increase in R&D spending was seen for all three different sized firms, while the increase in employment was mostly limited to larger firms. The widespread effects on R&D spending indicated that the state R&D tax credits had worked effectively; in other words, they encouraged most private firms to increase their R&D spending. The recipients ranged from the high-technology industry to the manufacturing and service industries. In particular, the state R&D tax credits also promoted the increase in R&D spending for small and medium firms as well as large firms.

8.2 POLICY IMPLICATIONS

In general, the findings presented here support the utilization of state R&D tax credits. As an indirect government intervention, state R&D tax credits could increase private R&D activities in two ways—by providing an incentive for encouraging additional activities *or* by providing a windfall. The possible windfall effect implies that the increase in private R&D activities could be achieved without credits. Judging whether an increase in presumed outcomes is newly created or merely a windfall is a fundamental question of every program evaluation. It is not an easy question to answer on the grounds that we cannot observe what would have happened without the program. Regarding this issue, the previous studies did not provide any reliable empirical evidence for the effects of state R&D tax credits, especially in a quantitative way. As a response to this fundamental question, this study insisted that there was an additional increase in R&D spending and employment resulting from the credits, which is separate from any increase from other relevant factors. In other words, an additional increase of outcomes from the credits could not be achieved without the credits, which tells us that the state R&D tax credits create new demands, not just windfalls. These findings were obtained by using multiple quantitative methods and multiple levels of analysis. In this sense, this study provided a fundamental piece of information about this specific policy tool for encouraging private R&D activities.

State R&D tax credits primarily aim to generate the *additional private R&D spending*, instead of providing government own funds and resources, and importantly this additional spending is supposed to be not able to realized without program as discussed above. Under the fundamental R&D-relevant assumptions, which are the private sector tends to underinvest R&D due to the possible exploitations and therefore, the actual level of private R&D is much lower than the desirable level of R&D for the society, government intervention for increasing the additional private R&D spending can result in increasing the benefits for the whole society. The additional R&D spending generated from state R&D tax credit programs is more meaningful for the society on the ground that the possible beneficiaries are not limited to the firms that directly received credits but extended to the firms that indirectly get a benefit through spillover effects. The importance of supporting private R&D can also be found as a primary generator of innovation and a source of increasing productivity, which are assumed to result in economic growth in the long run. Accordingly, the benefits from the state R&D tax credits can be spread out to the whole society through improving the overall level of productivity and eventually changing the economic system, not just limited to the firms received the credits or the high-technology industry directly and immediately. In this context, the state R&D tax credit program is a valuable social policy which generate a broader benefit for the whole society.

Meanwhile, as an indirect government intervention, state R&D tax credit programs provide flexibility and autonomy for private firms as follows: private firms can choose R&D projects freely depending on their needs and they can adjust the amount of R&D expenditure and number of employees based on their own decisions. In addition, most state R&D tax credits, including those evaluated in this study, are available to any private firm, based on qualified R&D expenditures over the base amount previously spent. These state R&D tax credit programs are generally applied to any individual firm within any industry.[121] However, as an indirect government intervention, some negative outcomes, such as the possibility of inclining short-term projects, are also expected. For achieving the higher social benefits and the higher spillover effects from the increased R&D, the balanced development of direct and indirect government interventions are required. Thus, complementary utilizations of indirect interventions with direct interventions are recommended for better supporting private R&D.

As a regional economic development program, state R&D tax credit programs can make a positive contribution to regional economic growth as well. First of all, this study found that

[121] Some state R&D tax credits are limited to specific industries; however, this type of tax credit is utilized in few states. These states are excluded in empirical analysis of this study.

there was an increase in high-technology industry employment. Obviously, increasing employment is one of the fundamental and explicit goals of every regional economic development policy. Encouraging private R&D is also recommended based on various regional economic growth theories. For example, according to economic base theory, R&D activities are positively associated with making new products that are exportable to other regions. More specifically, product cycle theory suggests that R&D activities are associated with the development of new products that possess some privileges before maturing and standardizing within cyclical product development. From flexible production theory, R&D also plays an important role in producing customized goods or niche market products through flexible specialization production systems. These regional economic growth theories also indicate that possible economic growth from the high-technology industry highly depends on the regional capacity including the level of regional industrial mix. Accordingly, the benefits from the state R&D tax credits could also highly depend on the level of industrial mix and more specifically could be limited to the region having a higher portion of high-technology industry or R&D-intensive industry. In sum, state R&D tax credits are also meaningful as a state-level or local-level policy in the sense that R&D activities and the high-technology industry are considered as a primary source of producing regional economic outcomes having a comparative advantage and consequently resulting in the differentiated regional economic growth. The empirical findings of this research is that state R&D tax credits have a positive effect on increases in R&D spending and the high-technology employees and these findings indicates positive effects from the program at the state level.

As I examined in Chapter 4, Section 4.3., each state has implemented the state R&D tax credit in quite different format. Namely, each state has developed its own state R&D tax credit program under its unique economic contexts. Accordingly, the revealed program effects should be interpreted within these contexts. For example, the state R&D tax credits in California and Massachusettes, having a well-developed program itself with the highest private R&D performance in the US, are expected to generate the larger benefits than any other state R&D tax credits. In addition, the state R&D tax credit in Indiana and Iowa, having a similar program but the lower level of private R&D performance, are expected to generate the positive but the limited program effect size.

In addition, even though some states with high private R&D performance did not include the treatment group in this study for contructing better comparison, those states recently adopted state R&D tax credits as well. Because this research was implemented by using multiple methods and multiple estimates with general categorizations for increasing external validity, the revealed program effects could be generalized to those states having this program recently. For example, Ohio and Texas, utilizing this program in a similar way with the treatment states in this research, are expected to have positive program effects within their areas. On the other hand, Michigan, one of the highest private R&D performance states, started to utilize this program from 2002 with quite limited operation by allowing the credit only to the pharmaceutical industry, and therefore the quited limited program effect is expected in this specific state. Evaluating the programs including those states could be an interesting further research topic.

As we know, the performance of private R&D acitivities in the US has been quite skewed and state R&D tax credits are eligible on the basis of the amount of R&D expenditures, we can expect that the larger private R&D expenditures can generate the larger increase in the amount of R&D expenditures by receiving the tax credits. It implies that the skewness of private R&D spending/activities cannot be adjusted or reduced through the state R&D tax credits. It could evoke the question why the government should support the private R&D within a state already reaching a high level of private R&D performance or whether state R&D tax credits is an appropriate policy within that state. We can answer this question within the context of the overall level of the private R&D for a society. Namely, encouraging the additional increase in private R&D spending even within states having a high level of private R&D performance is still significant on the basis of that the practical level of private R&D is much lower than the desirable level for a society. In addition, encouraging the additional increase in private R&D spending within states having a low level of private R&D performance is also important on the basis of that the gap between the practical level of private R&D and the desirable level for a society is expected to be much larger in those states.

Another policy implication can be made based on the analysis of the effects by firm size. This research found that state R&D tax credits were effective for generating a positive increase in R&D spending, regardless of firm size. Contrary to the common belief that small firms do not usually benefit from state R&D tax credits, this study found that small firms also received positive benefits for increasing their R&D spending from the credits. It is important to note that the revealed benefits included both direct and indirect effects; so the benefits to small firms could be indirect. This suggests that state R&D tax credits could provide incentives to increase R&D spending whether firms receive the credits directly or indirectly through spillovers or catch-up efforts, and that both types of possible incentives from the credits can be applied to firms of all sizes. The benefits from the state R&D tax credits by firm size could also depend on the firm characteristics. Some firms could focus on researches for finding and generating new idea while some firms could focus on development processes for producing new goods or improving existing production lines. Depending on individual projects that each firm performs, the impact of the partial increase in R&D spending generated from the state R&D tax credits could be huge or inconsiderable. Accordingly, the benefits from the state R&D tax credits could be differentiated by these firm characteristics.

This research also found that state R&D tax credits increased employment, for larger firms but not for small firms. An additional policy should therefore be implemented to encourage employment growth for small firms. For example, state R&D tax credits could be combined with a program specifically targeting small firms, such as SBIR or STTR, or a network could be established to help small firms find the appropriate human capital. A program specifically aimed at and focused on increasing employment for small firms is required in the sense that increasing employment is much harder than increasing spending, especially for small firms.

For better policy implementation, understanding the inner mechanisms of utilizing tax credits can provide rich information. For example, one relevant topic to examine is whether small firms and large firms utilize the credits in different ways based on their different roles in R&D activity. The way firms adjust their R&D investments and manpower plans in order to receive tax credits is another interesting topic for providing insight into better usages of tax credits. In addition, the effects on R&D spending and employment can be assessed for specific industries or specific firms. The effects also can be assessed for a specific type of employment and R&D spending. For example, if we can determine what types of employment are increased through the effects of the credits (e.g., basic research relevant employment, product development relevant employment, or sales relevant employment) the effects of the credits can be better understood and the policy can be better utilized. This kind of information could be obtained by in-depth case study or by the analysis of detailed datasets.

In sum, state R&D tax credits can be utilized for generating R&D spending and employment for future economic growth at the regional level. The increase in R&D could generate much larger benefits to the whole region than to specific firms due to the possibility of localized knowledge spillovers. It is also expected that this policy could increase R&D spending broadly for firms of all sizes while increasing employment on a limited basis for larger firms. The efficient utilization of this policy program, as an indirect government intervention, is also expected with complementary utilizations of direct funding or government-performed R&D projects.

8.3 CONTRIBUTIONS AND LIMITATIONS

8.3.1 Major contributions

This research contributed to the evaluation of public policy by developing evaluation strategies and quasi-experimental evaluation methods in particular. Based on the difficulty of evaluating regional economic development strategies due to unclear scopes of time, area, and beneficiaries, it was frequently mentioned that only few efforts were made to evaluate them. On the other hand, the lack of evaluation of R&D-relevant policies, especially quantitative evaluation, was also widely recognized. The evaluation strategies developed in this study could be generalized to evaluate regional level programs. As a response of prompt necessity for evaluating regional economic development policies, the advanced evaluation techniques could fill a gap between a theory and its realization. Furthermore, this study also provided variations of the possible applications of each method. This methodological development is one of the major contributions of this study.

In detail, the evaluation strategies developed in this study encompassed 1) how to define the hypothesis and rival hypotheses, 2) how to utilize multiple methods and multiple datasets for reliable and concrete analysis, 3) how to construct a valid treatment group and control group, and 4) how to control and capture the possible relevant effects from the relevant factors. First, multiple methods and multiple datasets can deal with multiple rival hypotheses and provide multiple findings with a complementary relationship. Second, a valid comparison group makes it possible to obtain more accurate and valid results for assessing program effects. Third, controlling other relevant factors is another key for better analysis and allows complementary relationships with major outcome changes to be detected.

An important issue for applying these techniques is data availability. Namely, the feasibility of performing program evaluations with these suggested methods depends largely on the possibility of utilizing appropriate datasets. In this sense, this study provided a major source of datasets that can be utilized for evaluating other economic development programs. Most of the datasets used in this study are easily accessible.

Next, empirical applications of the developed evaluation strategies are also meaningful for providing empirical evidence for the effectiveness of the selected policy and consequently supporting that policy. There are on-going debates on the effectiveness of R&D tax credit programs in terms of their necessity, permanent application, appropriate application range, and the boundary of their beneficiaries. This study could not answer all of these questions; however, it does at least answer the most basic question, which regards the necessity of tax credit utilization and has been ambiguous until now.

8.3.2 Limitations of this study

The methods utilized in this paper have some shortcomings. First, the analytical methods applied in this study could not recognize the sensitivity of the credit rate or the differences between the various states' credit laws because the program effects are estimated by the yearly average effect across states and industries. Accordingly, the revealed effects cannot reflect the possible difference by regional characteristics (i.e. R&D-intensive regions or less-relevant-to-R&D regions) and by firm characteristics for performing R&D activities (i.e. research-intensive firms, or development-intensive firms). However, this weakness could provide another advantage for analyzing the effectiveness of state R&D tax credits by treating each state law as a whole without incorporating any individual differences within the empirical specifications. This made it possible to avoid developing complex models, which was impractical in some senses.

Second, the results were somewhat the limited by the available datasets. This study could not estimate direct effects and indirect effects separately. With an extended dataset, including information on the recipients of tax credits, it would be possible to estimate direct effects and indirect effects separately. Separate estimates of direct and indirect effects could provide valuable information that could never be obtained without data on the recipients of credits. Next, with more extensive data for the high-technology industry, such as R&D spending by industry and state, better evidence of the effects for the high-technology industry could be provided. The relatively weak evidence of increased R&D spending for the high-technology industry could have resulted from the lack of possible datasets available. Again, the DD/DDD methods and the matching methods as quasi-experimental approaches require a huge amount of data. With larger datasets, plenty of information for improving programs could be provided.

8.4 CONCLUDING REMARKS

This study primarily aimed to develop an evaluation strategy for a regional level R&D policy, the state R&D tax credit. To do this, I used the quasi- experimental approach. The outcome changes between states with tax credits and states without tax credits were compared using evaluation strategies for dealing with the nonequivalent aspects of the two comparison groups.

As mentioned above in the description of research design (Chapter 6), there were allowances for existing plausible rival hypotheses. Indeed, some rival hypotheses were recognized, but could not be appropriately ruled out. Based on the selected methods, they can be ruled out indirectly as unobserved parts or unexplained parts within the empirical models. However, there were still allowances for finding the better measurement. It all depends on current knowledge and context. Therefore, the findings reported here should be understood to be true based only on the current knowledge and context and should be considered tentative. This is the most important assumption of quasi-experimental approaches, which we must keep in mind.

APPENDIX A

THE DETAILED STATISTICAL RESULTS

The appendix A includes the detailed statistical results for the analysis of effects of state R&D tax credits. They are 1) the estimates of effects on R&D spending, 2) the estimates of effects on employment, 3) the estimates of effects on R&D spending by firm size, and 4) the estimates of effects on employment by firm size. Under each category, the DD/DDD estimates and the matching estimates and those for the whole industry and the high technology industry are presented. There are both the state and firm level analyses for the first two categories as the estimates of overall effects and only the firm level analysis for the last two categories as disaggregated analyses by firm size (defined as small, medium, and large). There are also both the state and firm level analyses for the DD/DDD estimates and only the firm level analysis for the matching estimates.

A.1 THE DETAILED STATISTICAL RESULTS FOR THE ANALYSIS OF R&D SPENDING

A.1.1 The DD estimates for the all industry (the state and firm levels)

DD estimates with covariates (GSP), at the state level, in level

```
Regression with robust standard errors                Number of obs =     185
                                                       F( 18,   137) = 3221.56
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9819
                                                       Adj R-squared =  0.9757
                                                       Root MSE      =  480.11
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
rdprivat_con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   628.9734   189.0939     3.33   0.001     255.0532    1002.894
  laborforce |   .0002633   .0003034     0.87   0.387    -.0003366    .0008632
        know |   .0830336   .0515678     1.61   0.110    -.0189381    .1850053
    unemprte |  -91.64896    38.4378    -2.38   0.018    -167.6571   -15.64085
      laggsp |   .0052068   .0035398     1.47   0.144    -.0017929    .0122065
    lagrdpri |   .8181541   .1016521     8.05   0.000     .6171441    1.019164
    lagrdfed |  -.1533866   .1092177    -1.40   0.162     -.369357    .0625839
------------------------------------------------------------------------------
    fips |   absorbed                                        (29 categories)
```

DD estimates with covariates (personal income), at the state level, in level

```
Regression with robust standard errors                Number of obs =     239
                                                       F( 21,   186) =18433.18
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9803
                                                       Adj R-squared =  0.9748
                                                       Root MSE      =  445.27
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
rdprivat_con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   404.5683   125.4084     3.23   0.001     157.1625    651.9741
  laborforce |   .0003866   .0002045     1.89   0.060    -.0000169    .0007901
        know |   .0057643   .0331739     0.17   0.862    -.0596812    .0712098
    unemprte |    -63.815   36.32649    -1.76   0.081    -135.4799      7.8499
lagperincome |  -1.93e-06   4.65e-06    -0.41   0.679    -.0000111    7.25e-06
    lagrdpri |   .9498772   .0871857    10.89   0.000     .7778772    1.121877
    lagrdfed |  -.1040123   .0928594    -1.12   0.264    -.2872054    .0791807
------------------------------------------------------------------------------
    fips |   absorbed                                        (30 categories)
```

DD estimates with covariates, at the firm level, in level

```
Regression with robust standard errors                 Number of obs =   36977
                                                       F( 34, 36774) =10413.84
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9689
                                                       Adj R-squared =  0.9687
                                                       Root MSE      =   27.86
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      rd_con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   1.982597   .8180954     2.42   0.015     .3791072    3.586088
    lagsales |   .0021369   .0012529     1.71   0.088    -.0003188    .0045926
  knowcounty |   .0002173   .0001647     1.32   0.187    -.0001055    .0005401
   emp2digit |   1.53e-06   7.27e-07     2.10   0.036     1.03e-07    2.95e-06
    unemprte |  -.1200387   .0789825    -1.52   0.129    -.2748467    .0347694
       lagrd |   .9735691   .0328358    29.65   0.000       .90921    1.037928
------------------------------------------------------------------------------
   fips |   absorbed                                       (36 categories)
```

DD estimates with covariates (GSP), at the state level, in log

```
Regression with robust standard errors                 Number of obs =     185
                                                       F( 18,   137) =  441.99
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9746
                                                       Adj R-squared =  0.9659
                                                       Root MSE      =  .21738
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
    logrdpri |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   .1076201   .0956295     1.13   0.262    -.0814806    .2967208
    loglabor |  -.7148014   .6207164    -1.15   0.252    -1.942225    .5126226
     logknow |   .4270757   .2583994     1.65   0.101    -.0838915    .9380428
    logunemp |  -.0944159   .0988541    -0.96   0.341    -.2898932    .1010613
   loglaggsp |   .6841605   .5133205     1.33   0.185    -.3308954    1.699216
 loglagrdpri |   .4329663    .110532     3.92   0.000     .2143969    .6515357
 loglagrdfed |  -.0340026   .0483575    -0.70   0.483    -.1296262    .0616211
------------------------------------------------------------------------------
    fips |   absorbed                                       (29 categories)
```

DD estimates with covariates (personal income), at the state level, in log

```
Regression with robust standard errors                 Number of obs =     239
                                                       F( 21,   186) =  713.23
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9736
                                                       Adj R-squared =  0.9662
                                                       Root MSE      =  .21771
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
    logrdpri |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   .1384532   .0771181     1.80   0.074    -.0136854    .2905918
    loglabor |   -.189066   .5682592    -0.33   0.740    -1.310128    .9319959
     logknow |   .4258332   .2788024     1.53   0.128    -.1241882    .9758546
    logunemp |  -.1266722   .0955555    -1.33   0.187    -.3151841    .0618398
loglagperi~e |   .3386492    .507619     0.67   0.506    -.6627817     1.34008
 loglagrdpri |   .5021876   .0875922     5.73   0.000     .3293858    .6749895
```

```
 loglagrdfed |  -.0445006   .0322299    -1.38   0.169    -.1080837    .0190825
------------------------------------------------------------------------------
        fips |   absorbed                                     (30 categories)
```

DD estimates with covariates, at the firm level, in log

```
Regression with robust standard errors                 Number of obs =   33564
                                                       F( 34, 33370) =12287.53
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9399
                                                       Adj R-squared =  0.9396
                                                       Root MSE      =  .52968
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   .0452692   .0130774     3.46   0.001      .019637    .0709014
 loglagsales |   .0467196   .0057684     8.10   0.000     .0354133    .0580259
logknowcou~y |   .0058641   .0018716     3.13   0.002     .0021957    .0095325
   logemp2di |   .0092236   .0054934     1.68   0.093    -.0015437    .0199909
    logunemp |   .0206178   .0181481     1.14   0.256    -.0149532    .0561887
    loglagrd |   .9288579   .0061525   150.97   0.000     .9167988     .940917
------------------------------------------------------------------------------
        fips |   absorbed                                     (36 categories)
```

A.1.2 The matching estimates for the all industry (the firm level)

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched |  18.091704   20.0823125  -1.99060845
                        ATT |  18.091704   6.36071447   11.7309896
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       384 |       384
   Treated |       753 |       753
-----------+-----------+----------
     Total |     1,137 |     1,137
Bootstrap statistics                              Number of obs    =      1411
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  11.73099  .9049056   3.28729   5.124933   18.33705   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1, in level

```
Probit estimates                                  Number of obs   =       1137
                                                  LR chi2(15)     =     697.71
                                                  Prob > chi2     =     0.0000
Log likelihood =  -378.2762                       Pseudo R2       =     0.4798
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |   .0012731   .0007816     1.63   0.103    -.0002589     .002805
    lagsales |  -.0000902   .0000667    -1.35   0.176    -.0002209    .0000404
  knowcounty |   .0003687   .0000555     6.65   0.000     .0002599    .0004774
    emp2digit |  -2.11e-07   5.45e-07    -0.39   0.698    -1.28e-06    8.56e-07
```

```
 firmsizeall |   .0015063   .0050921     0.30   0.767    -.0084741    .0114866
incapita89~n |  -.0000307   .0000468    -0.66   0.512    -.0001224     .000061
     edu1290 |   .0079441   .0193302     0.41   0.681    -.0299424    .0458306
     edu1690 |  -.0663324     .01946    -3.41   0.001    -.1044732   -.0281915
    sbir8690 |   8.40e-08   5.46e-09    15.37   0.000     7.33e-08    9.47e-08
    univ8690 |  -3.53e-06   2.95e-07   -11.96   0.000    -4.11e-06   -2.95e-06
   popchange |  -2.297268   .3798961    -6.05   0.000    -3.041851   -1.552685
     crime91 |    .000115   .0000272     4.23   0.000     .0000617    .0001683
   poverty89 |  -.1166838   .0306984    -3.80   0.000    -.1768516   -.0565161
  unemprat90 |  -.3137306   .0704192    -4.46   0.000    -.4517497   -.1757114
mestdensi~87 |   404.7048   97.80679     4.14   0.000     213.0071    596.4026
       _cons |   3.536053   1.694904     2.09   0.037     .2141022    6.858004
-----------------------------------------------------------------
-----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
        rd_con_d  Unmatched |  18.091704   20.0823125  -1.99060845
                        ATT | 21.4422841   15.2535778   6.18870626
----------------------------+-------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        384 |       384
   Treated |       504        249 |       753
-----------+----------------------+----------
     Total |       504        633 |     1,137

Bootstrap statistics                               Number of obs    =      1137
                                                   Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  6.188706  2.593112  13.66744  -21.27704   33.65446   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3, in level

```
-----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
        rd_con_d  Unmatched |  18.091704   20.0823125  -1.99060845
                        ATT | 21.4422841   15.9219213   5.52036281
----------------------------+-------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        384 |       384
   Treated |       504        249 |       753
-----------+----------------------+----------
     Total |       504        633 |     1,137
Bootstrap statistics                               Number of obs    =      1137
                                                   Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  5.520363  .8824183  8.636435   -11.8352   22.87593   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log

```
-----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
          logrd_d  Unmatched |  .50242481   .255669252   .246755557
                        ATT |  .50242481   .456125396   .046299414
----------------------------+-------------------------------------
```

```
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       384 |       384
   Treated |       753 |       753
-----------+-----------+----------
     Total |     1,137 |     1,137

Bootstrap statistics                                  Number of obs    =      1411
                                                      Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0462994  .0522673  .1603999  -.2760363   .3686352   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1, in log

```
Probit estimates                                      Number of obs   =       1137
                                                      LR chi2(15)     =     697.71
                                                      Prob > chi2     =     0.0000
Log likelihood =  -378.2762                           Pseudo R2       =     0.4798

------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |   .0012731   .0007816     1.63   0.103    -.0002589     .002805
    lagsales |  -.0000902   .0000667    -1.35   0.176    -.0002209    .0000404
  knowcounty |   .0003687   .0000555     6.65   0.000     .0002599    .0004774
   emp2digit |  -2.11e-07   5.45e-07    -0.39   0.698    -1.28e-06    8.56e-07
 firmsizeall |   .0015063   .0050921     0.30   0.767    -.0084741    .0114866
incapita89~n |  -.0000307   .0000468    -0.66   0.512    -.0001224     .000061
     edu1290 |   .0079441   .0193302     0.41   0.681    -.0299424    .0458306
     edu1690 |  -.0663324     .01946    -3.41   0.001    -.1044732   -.0281915
    sbir8690 |   8.40e-08   5.46e-09    15.37   0.000     7.33e-08    9.47e-08
    univ8690 |  -3.53e-06   2.95e-07   -11.96   0.000    -4.11e-06   -2.95e-06
   popchange |  -2.297268   .3798961    -6.05   0.000    -3.041851   -1.552685
     crime91 |    .000115   .0000272     4.23   0.000     .0000617    .0001683
   poverty89 |  -.1166838   .0306984    -3.80   0.000    -.1768516   -.0565161
  unemprat90 |  -.3137306   .0704192    -4.46   0.000    -.4517497   -.1757114
mestdensi~87 |   404.7048   97.80679     4.14   0.000     213.0071    596.4026
       _cons |   3.536053   1.694904     2.09   0.037     .2141022    6.858004
------------------------------------------------------------------------------
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
         logrd_d  Unmatched |  .50242481   .255669252   .246755557
                        ATT | .458672774   .395386236   .063286538
----------------------------+-----------------------------------

 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        384 |       384
   Treated |       403        350 |       753
-----------+----------------------+----------
     Total |       403        734 |     1,137

Bootstrap statistics                                  Number of obs    =      1137
                                                      Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0632865  .0846884  .1455931  -.2292937   .3558668   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3, in log

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------------------
         logrd_d  Unmatched |  .50242481   .255669252   .246755557
                        ATT | .458672774   .366315996   .092356778
----------------------------+-----------------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        384 |       384
   Treated |       403        350 |       753
-----------+----------------------+----------
     Total |       403        734 |     1,137

Bootstrap statistics                              Number of obs    =      1137
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0923568  .0426357  .1126391  -.1339999   .3187135   (N)
------------------------------------------------------------------------------
```

A.1.3 The DD/DDD estimates for the high-technology industry (the firm level)

DD estimates with covariates, at the firm level, in level

```
Regression with robust standard errors                 Number of obs =   22373
                                                       F( 30, 22284) =98831.36
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9694
                                                       Adj R-squared =  0.9692
                                                       Root MSE      =  34.227
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      rd_con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   2.859282   1.343354     2.13   0.033     .2262132     5.49235
    lagsales |   .0025341   .0014932     1.70   0.090    -.0003927    .0054608
  knowcounty |   .0004222   .0002157     1.96   0.050    -6.26e-07    .0008451
   emp2digit |   1.26e-06   7.86e-07     1.61   0.108    -2.79e-07    2.80e-06
     unemprte |  -.1527327    .129455    -1.18   0.238    -.4064735    .1010082
       lagrd |   .9683011   .0380158    25.47   0.000     .8937875    1.042815
------------------------------------------------------------------------------
    fips |   absorbed                                        (36 categories)
```

DDD estimates with covariates, at the firm level, in level

```
Regression with robust standard errors                 Number of obs =   36977
                                                       F( 34, 36771) =56463.52
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9689
                                                       Adj R-squared =  0.9687
                                                       Root MSE      =   27.86
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      rd_con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostddd |   1.907444   1.276049     1.49   0.135     -.593649    4.408537
```

```
    lagsales |   .0021368   .0012522     1.71   0.088    -.0003177    .0045912
  knowcounty |   .0002144   .0001616     1.33   0.185    -.0001024    .0005312
   emp2digit |   1.50e-06   7.43e-07     2.02   0.044     4.22e-08    2.96e-06
    unemprte |   -.121922   .0772591    -1.58   0.115    -.2733521    .0295082
       lagrd |   .9735862   .0328186    29.67   0.000     .9092608    1.037912
------------------------------------------------------------------------------
   fips |   absorbed                                     (36 categories)
```

DD estimates with covariates, at the firm level, in level

```
Regression with robust standard errors                 Number of obs =   20557
                                                       F( 30, 20468) =  738.72
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9387
                                                       Adj R-squared =  0.9384
                                                       Root MSE      =  .53711
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .0200786   .0203866     0.98   0.325    -.0198808    .0600381
 loglagsales |   .0380972   .0065461     5.82   0.000     .0252663     .050928
logknowcou~y |   .0090493   .0029643     3.05   0.002      .003239    .0148595
   logemp2di |   .0193415   .0068031     2.84   0.004     .0060069    .0326761
    logunemp |    .026503   .0277331     0.96   0.339    -.0278561    .0808622
    loglagrd |   .9320037   .0075732   123.07   0.000     .9171596    .9468478
------------------------------------------------------------------------------
   fips |   absorbed                                     (36 categories)
```

DDD estimates with covariates, at the firm level, in log

```
Regression with robust standard errors                 Number of obs =   33564
                                                       F( 34, 33367) = 6613.50
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9399
                                                       Adj R-squared =  0.9396
                                                       Root MSE      =  .52964
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   tpostddd |  -.0361532   .0272598    -1.33   0.185    -.0895833    .0172768
 loglagsales |   .0470313   .0057317     8.21   0.000      .035797    .0582656
logknowcou~y |   .0058848   .0018675     3.15   0.002     .0022244    .0095451
   logemp2di |    .008471   .0054289     1.56   0.119    -.0021699    .0191119
    logunemp |   .0185529   .0183237     1.01   0.311    -.0173622     .054468
    loglagrd |   .9285715   .0060939   152.38   0.000     .9166272    .9405157
------------------------------------------------------------------------------
   fips |   absorbed                                     (36 categories)
```

A.1.4 The matching estimates for the high technology industry (the firm level)

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | 23.4038383   34.0628364  -10.6589981
                        ATT | 23.4038383   4.99466947   18.4091689
----------------------------+-----------------------------------
```

```
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       216 |       216
   Treated |       509 |       509
-----------+-----------+----------
     Total |       725 |       725

Bootstrap statistics                              Number of obs    =      1411
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  18.40917  -.362499  4.387841   9.591472   27.22686   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1, in level

```
Probit estimates                                  Number of obs   =        725
                                                  LR chi2(15)     =     471.31
                                                  Prob > chi2     =     0.0000
Log likelihood = -205.94183                       Pseudo R2       =     0.5336
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |   .0010815   .0010054     1.08   0.282    -.0008889     .003052
    lagsales |  -.0001244   .0000913    -1.36   0.173    -.0003034    .0000546
  knowcounty |   .0002656   .0000764     3.48   0.001     .0001158    .0004153
   emp2digit |  -3.53e-07   6.55e-07    -0.54   0.590    -1.64e-06    9.32e-07
 firmsizeall |    .008175   .0126156     0.65   0.517    -.0165513    .0329012
incapita89~n |  -.0000156   .0000612    -0.25   0.799    -.0001356    .0001044
     edu1290 |   .0023621   .0263056     0.09   0.928    -.0491959    .0539202
     edu1690 |  -.0755997   .0248056    -3.05   0.002    -.1242178   -.0269816
    sbir8690 |   8.55e-08   7.01e-09    12.21   0.000     7.18e-08    9.92e-08
    univ8690 |  -3.64e-06   3.83e-07    -9.51   0.000    -4.39e-06   -2.89e-06
   popchange |  -1.961652   .4376214    -4.48   0.000    -2.819374   -1.103929
     crime91 |   .0000462   .0000383     1.21   0.227    -.0000288    .0001212
   poverty89 |  -.0799225   .0401045    -1.99   0.046    -.1585259   -.0013191
  unemprat90 |  -.2768549   .0961748    -2.88   0.004    -.4653541   -.0883556
mestdensi~87 |   472.5369    136.169     3.47   0.001     205.6506    739.4232
       _cons |   3.890507   2.411673     1.61   0.107    -.8362857      8.6173
------------------------------------------------------------------------------
----------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------------
        rd_con_d  Unmatched | 23.4038383   34.0628364  -10.6589981
                        ATT | 30.6070301   17.4719607   13.1350694
----------------------------+-----------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        216 |       216
   Treated |       342        167 |       509
-----------+----------------------+----------
     Total |       342        383 |       725

Bootstrap statistics                              Number of obs    =       725
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  13.13507 -9.379962  18.74574   -24.5359   50.80604   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
        rd_con_d  Unmatched | 23.4038383   34.0628364  -10.6589981
                        ATT | 30.6070301   13.0186017   17.5884284
----------------------------+-------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        216 |       216
   Treated |       342        167 |       509
-----------+----------------------+----------
     Total |       342        383 |       725

Bootstrap statistics                              Number of obs    =       725
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  17.58843 -12.09015  16.41963  -15.40805   50.58491   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
         logrd_d  Unmatched | .613503569   .395233167   .218270402
                        ATT | .613503569   .180464386   .433039183
----------------------------+-------------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       216 |       216
   Treated |       509 |       509
-----------+-----------+----------
     Total |       725 |       725

Bootstrap statistics                              Number of obs    =      1411
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .4330392 -.0739949  .3217609  -.2135636   1.079642   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1, in log

```
Probit estimates                                  Number of obs   =        725
                                                  LR chi2(15)     =     471.31
                                                  Prob > chi2     =     0.0000
Log likelihood = -205.94183                       Pseudo R2       =     0.5336
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |   .0010815   .0010054     1.08   0.282    -.0008889     .003052
    lagsales |  -.0001244   .0000913    -1.36   0.173    -.0003034    .0000546
  knowcounty |   .0002656   .0000764     3.48   0.001     .0001158    .0004153
   emp2digit |  -3.53e-07   6.55e-07    -0.54   0.590    -1.64e-06    9.32e-07
 firmsizeall |    .008175   .0126156     0.65   0.517    -.0165513    .0329012
incapita89~n |  -.0000156   .0000612    -0.25   0.799    -.0001356    .0001044
     edu1290 |   .0023621   .0263056     0.09   0.928    -.0491959    .0539202
     edu1690 |  -.0755997   .0248056    -3.05   0.002    -.1242178   -.0269816
    sbir8690 |   8.55e-08   7.01e-09    12.21   0.000     7.18e-08    9.92e-08
    univ8690 |  -3.64e-06   3.83e-07    -9.51   0.000    -4.39e-06   -2.89e-06
```

```
   popchange |  -1.961652   .4376214    -4.48   0.000    -2.819374   -1.103929
     crime91 |   .0000462   .0000383     1.21   0.227    -.0000288    .0001212
   poverty89 |  -.0799225   .0401045    -1.99   0.046    -.1585259   -.0013191
  unemprat90 |  -.2768549   .0961748    -2.88   0.004    -.4653541   -.0883556
mestdensi~87 |   472.5369    136.169     3.47   0.001     205.6506    739.4232
       _cons |   3.890507   2.411673     1.61   0.107    -.8362857      8.6173
------------------------------------------------------------------------------
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
         logrd_d  Unmatched | .613503569   .395233167   .218270402
                        ATT | .544351209   .483040293   .061310916
----------------------------+-------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        216 |       216
   Treated |       401        108 |       509
-----------+----------------------+----------
     Total |       401        324 |       725

Bootstrap statistics                              Number of obs    =       725
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0613109  .1238898   .251484  -.4440651    .566687   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3, in log

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
         logrd_d  Unmatched | .613503569   .395233167   .218270402
                        ATT | .544351209   .487848385   .056502824
----------------------------+-------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        216 |       216
   Treated |       401        108 |       509
-----------+----------------------+----------
    Total |       401        324 |       725

Bootstrap statistics                              Number of obs    =       725
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0565028  .0968315  .2554959  -.4569354    .569941   (N)
------------------------------------------------------------------------------
```

A.2 THE DETAILED STATISTICAL RESULTS FOR THE ANALYSIS FOR EMPLOYMENT

A.2.1 The DD estimates for the all industry (the state and firm levels)

DD estimates with covariates (GSP), at the state level, in level

```
Regression with robust standard errors                Number of obs =     355
                                                      F( 18,   307) =32051.99
                                                      Prob > F      =  0.0000
                                                      R-squared     =  0.9974
                                                      Adj R-squared =  0.9971
                                                      Root MSE      =  20.035
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   2.366136   6.105525     0.39   0.699    -9.647835    14.38011
  laborforce |   5.19e-06   5.50e-06     0.94   0.346    -5.64e-06     .000016
      laggsp |   7.17e-06   .0000967     0.07   0.941    -.0001831    .0001975
        know |  -.0004972   .0019416    -0.26   0.798    -.0043177    .0033233
    lagrdpri |   .0006074   .0019296     0.31   0.753    -.0031895    .0044043
    lagrdfed |   .0011814   .0023877     0.49   0.621    -.0035169    .0058796
      lagemp |   1.009409   .0074981   134.62   0.000     .9946552    1.024164
------------------------------------------------------------------------------
    fips |   absorbed                                       (29 categories)
```

DD estimates with covariates (personal income), at the state level, in level

```
Regression with robust standard errors                Number of obs =     450
                                                      F( 22,   398) = 1.7e+06
                                                      Prob > F      =  0.0000
                                                      R-squared     =  0.9967
                                                      Adj R-squared =  0.9963
                                                      Root MSE      =  21.836
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   4.939351   4.614423     1.07   0.285    -4.132339    14.01104
  laborforce |   3.96e-06   3.97e-06     1.00   0.318    -3.83e-06    .0000118
lagperincome |   2.09e-08   8.73e-08     0.24   0.811    -1.51e-07    1.92e-07
        know |   .0008204   .0011222     0.73   0.465    -.0013858    .0030266
    lagrdpri |  -.0000902   .0007774    -0.12   0.908    -.0016185     .001438
    lagrdfed |  -.0001249   .0018353    -0.07   0.946     -.003733    .0034832
      lagemp |   1.007988   .0083188   121.17   0.000     .9916342    1.024343
------------------------------------------------------------------------------
    fips |   absorbed                                       (29 categories)
```

DD estimates with covariate, at the firm level, in level

```
Regression with robust standard errors                Number of obs =   35435
                                                      F( 34, 35231) =51060.15
                                                      Prob > F      =  0.0000
                                                      R-squared     =  0.9855
```

```
                                                       Adj R-squared =  0.9854
                                                       Root MSE      =  3.5915
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |    .2045032   .1263976     1.62   0.106      -.04324    .4522465
       lagrd |  -.0010939   .0027821    -0.39   0.694    -.0065469    .0043592
      lagemp |   .9960088   .0205554    48.45   0.000     .9557196    1.036298
    lagsales |  -.0000506   .0001068    -0.47   0.635      -.00026    .0001587
  knowcounty |  -4.55e-06   1.00e-05    -0.46   0.649    -.0000241     .000015
   emp2digit |   6.78e-08   1.12e-07     0.61   0.544    -1.51e-07    2.87e-07
------------------------------------------------------------------------------
        fips |   absorbed                                     (36 categories)
```

DD estimates with covariates (GSP), at the state level, in log

```
Regression with robust standard errors                 Number of obs =     355
                                                       F( 18,   307) = 6314.61
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9940
                                                       Adj R-squared =  0.9931
                                                       Root MSE      =  .11914
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |    .0275245   .0391448     0.70   0.482    -.0495015    .1045505
    loglabor |   .3660345   .2383297     1.54   0.126    -.1029319    .8350009
   loglaggsp |  -.2553853   .1899592    -1.34   0.180    -.6291722    .1184015
     logknow |   .0358318   .0975597     0.37   0.714    -.1561384     .227802
 loglagrdpri |  -.0082126    .036007    -0.23   0.820    -.0790642    .0626391
 loglagrdfed |  -.0033356   .0048518    -0.69   0.492    -.0128827    .0062114
   loglagemp |   .9866757   .0143407    68.80   0.000     .9584573    1.014894
------------------------------------------------------------------------------
        fips |   absorbed                                     (29 categories)
```

DD estimates with covariates (personal income), at the state level, in log

```
Regression with robust standard errors                 Number of obs =     450
                                                       F( 22,   398) =17886.37
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9946
                                                       Adj R-squared =  0.9939
                                                       Root MSE      =  .11038

                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |    .0358537   .0281361     1.27   0.203    -.0194602    .0911676
    loglabor |   .3542803   .1877443     1.89   0.060    -.0148142    .7233748
loglagperi~e |  -.3736649   .2473335    -1.51   0.132    -.8599083    .1125784
     logknow |   .0910667   .0760016     1.20   0.232    -.0583481    .2404815
 loglagrdpri |  -.0316543   .0378575    -0.84   0.404    -.1060799    .0427713
 loglagrdfed |  -.0040858   .0047067    -0.87   0.386    -.0133389    .0051673
   loglagemp |   .9875559   .0123681    79.85   0.000     .9632411    1.011871
------------------------------------------------------------------------------
        fips |   absorbed                                     (29 categories)
```

DD estimates with covariates, at the firm level, in log

```
Regression with robust standard errors                 Number of obs =   32438
                                                       F( 34, 32243) =21649.53
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9779
                                                       Adj R-squared =  0.9777
                                                       Root MSE      =  .32956
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .0079976   .0108309     0.74   0.460    -.0132315    .0292266
 loglagsales |   .0293512   .0044629     6.58   0.000     .0206037    .0380988
logknowcou~y |  -.0023315   .0011684    -2.00   0.046    -.0046216   -.0000414
   logemp2di |   .0032413    .003204     1.01   0.312    -.0030387    .0095213
   loglagemp |   .9381875   .0056946   164.75   0.000     .9270258    .9493491
    loglagrd |    .019277   .0026393     7.30   0.000     .0141039      .02445
------------------------------------------------------------------------------
   fips |   absorbed                                      (36 categories)
```

A.2.2 The matching estimates for the all industry (the firm level)

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .612008183  -.422384231   1.03439241
                        ATT | .612008183  -.136910079   .748918262
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       419 |       419
   Treated |       778 |       778
-----------+-----------+----------
     Total |     1,197 |     1,197

Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .7489182 -.0580058  .5306412  -.3174451   1.815282   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level

```
Probit estimates                                  Number of obs   =       1197
                                                  LR chi2(16)     =     749.25
                                                  Prob > chi2     =     0.0000
Log likelihood = -400.39789                       Pseudo R2       =     0.4834
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |   .0159034   .0142873     1.11   0.266    -.0120993    .0439061
       lagrd |   .0010422   .0008414     1.24   0.216     -.000607    .0026914
    lagsales |  -.0001349   .0000789    -1.71   0.087    -.0002895    .0000197
  knowcounty |   .0003566   .0000531     6.72   0.000     .0002525    .0004607
   emp2digit |  -2.50e-07   5.23e-07    -0.48   0.633    -1.28e-06    7.75e-07
```

```
 firmsizeall |  -.0084981   .0113568    -0.75   0.454     -.030757    .0137609
incapita89~n |  -.0000445   .0000461    -0.97   0.334    -.0001348    .0000458
     edu1290 |   -.003766   .0182947    -0.21   0.837    -.0396229    .0320909
     edu1690 |   -.059405   .0189085    -3.14   0.002    -.0964651    -.022345
    sbir8690 |   8.32e-08   5.28e-09    15.76   0.000     7.28e-08    9.35e-08
    univ8690 |  -3.46e-06   2.86e-07   -12.08   0.000    -4.02e-06   -2.90e-06
   popchange |  -2.216854   .3712417    -5.97   0.000    -2.944474   -1.489234
     crime91 |   .0001143   .0000266     4.30   0.000     .0000621    .0001664
   poverty89 |  -.1399547   .0297231    -4.71   0.000    -.1982109   -.0816986
  unemprat90 |  -.2681675   .0675241    -3.97   0.000    -.4005124   -.1358226
mestdensi~87 |   456.0516   96.06086     4.75   0.000     267.7758    644.3275
       _cons |    4.31789   1.628571     2.65   0.008     1.125949    7.509831
------------------------------------------------------------------------------
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .612008183  -.422384231   1.03439241
                        ATT | .493632977   1.52002947  -1.02639649
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        419 |       419
   Treated |       395        383 |       778
-----------+----------------------+----------
     Total |       395        802 |     1,197
Bootstrap statistics                              Number of obs    =      1197
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -1.026397  1.526512  1.642524  -4.327172   2.274379   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3, in level

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .612008183  -.422384231   1.03439241
                        ATT | .493632977   1.25751287  -.763879893
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        419 |       419
   Treated |       395        383 |       778
-----------+----------------------+----------
     Total |       395        802 |     1,197

Bootstrap statistics                              Number of obs    =      1197
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -.7638799  1.545056  1.677197  -4.134332   2.606573   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
         logemp_d  Unmatched |   .2936104   .112359127   .181251273
                        ATT |   .2936104   .247785843   .045824556
----------------------------+-----------------------------------
```

```
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       419 |       419
   Treated |       778 |       778
-----------+-----------+----------
     Total |     1,197 |     1,197

Bootstrap statistics                                  Number of obs    =      1502
                                                      Replications     =        50
------------------------------------------------------------------------------
Variable      |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0458246  .0446429  .0967581  -.1486182   .2402673   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1, in log

```
Probit estimates                                      Number of obs   =       1197
                                                      LR chi2(16)     =     749.25
                                                      Prob > chi2     =     0.0000
Log likelihood = -400.39789                           Pseudo R2       =     0.4834
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |   .0159034   .0142873     1.11   0.266    -.0120993    .0439061
       lagrd |   .0010422   .0008414     1.24   0.216     -.000607    .0026914
    lagsales |  -.0001349   .0000789    -1.71   0.087    -.0002895    .0000197
  knowcounty |   .0003566   .0000531     6.72   0.000     .0002525    .0004607
   emp2digit |  -2.50e-07   5.23e-07    -0.48   0.633    -1.28e-06    7.75e-07
 firmsizeall |  -.0084981   .0113568    -0.75   0.454     -.030757    .0137609
incapita89~n |  -.0000445   .0000461    -0.97   0.334    -.0001348    .0000458
     edu1290 |   -.003766   .0182947    -0.21   0.837    -.0396229    .0320909
     edu1690 |   -.059405   .0189085    -3.14   0.002    -.0964651    -.022345
    sbir8690 |   8.32e-08   5.28e-09    15.76   0.000     7.28e-08    9.35e-08
    univ8690 |  -3.46e-06   2.86e-07   -12.08   0.000    -4.02e-06   -2.90e-06
   popchange |  -2.216854   .3712417    -5.97   0.000    -2.944474   -1.489234
     crime91 |   .0001143   .0000266     4.30   0.000     .0000621    .0001664
   poverty89 |  -.1399547   .0297231    -4.71   0.000    -.1982109   -.0816986
  unemprat90 |  -.2681675   .0675241    -3.97   0.000    -.4005124   -.1358226
mestdensi~87 |   456.0516   96.06086     4.75   0.000     267.7758    644.3275
       _cons |    4.31789   1.628571     2.65   0.008     1.125949    7.509831
------------------------------------------------------------------------------
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched |   .2936104   .112359127   .181251273
                        ATT | .295989497   .281235987    .01475351
----------------------------+-----------------------------------

 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        419 |       419
   Treated |       395        383 |       778
-----------+----------------------+----------
     Total |       395        802 |     1,197

Bootstrap statistics                                  Number of obs    =      1197
                                                      Replications     =        50
------------------------------------------------------------------------------
Variable      |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0147535  .0316939  .1552191   -.297171    .326678   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3, in log

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
        logemp_d  Unmatched |    .2936104   .112359127   .181251273
                        ATT | .295989497   .234488162   .061501335
----------------------------+-------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        419 |       419
   Treated |       395        383 |       778
-----------+----------------------+----------
     Total |       395        802 |     1,197

Bootstrap statistics                              Number of obs    =      1197
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0615013 -.0084342  .1644154  -.2689038   .3919065   (N)
------------------------------------------------------------------------------
```

A.2.3 The DD/DDD estimates for the high-technology industry (the state and firm levels)

DD estimates with covariates (GSP), at the state level, in level

```
Regression with robust standard errors                 Number of obs =     164
                                                       F( 18,   122) =31777.49
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9952
                                                       Adj R-squared =  0.9935
                                                       Root MSE      =   13.47
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |  -2.704118   4.511972    -0.60   0.550    -11.63602     6.22778
  laborforce |    .000012   8.97e-06     1.34   0.181    -5.70e-06    .0000298
      laggsp |  -.0001621   .0001054    -1.54   0.127    -.0003708    .0000465
        know |   -.001604   .0020324    -0.79   0.432    -.0056273    .0024194
    lagrdpri |   .0005346   .0015067     0.35   0.723    -.0024481    .0035173
    lagrdfed |  -.0034081   .0022011    -1.55   0.124    -.0077655    .0009493
      lagemp |   .9724642   .0283413    34.31   0.000     .9163597    1.028569
------------------------------------------------------------------------------
    fips |   absorbed                                       (23 categories)
```

DD estimates with covariates (personal income), at the state level, in level

```
Regression with robust standard errors                 Number of obs =     206
                                                       F( 20,   160) =19252.70
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9942
                                                       Adj R-squared =  0.9926
                                                       Root MSE      =  13.921
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
```

```
     tpostdd |   1.063255   3.581101     0.30   0.767    -6.009067    8.135576
  laborforce |   .0000105   5.26e-06     2.00   0.047     1.42e-07    .0000209
lagperincome |  -1.52e-07   1.03e-07    -1.47   0.142    -3.55e-07    5.14e-08
        know |  -.0009226   .0015699    -0.59   0.558    -.0040231    .0021779
    lagrdpri |  -.0003903   .0005711    -0.68   0.495     -.001518    .0007375
    lagrdfed |  -.0049662   .0018342    -2.71   0.008    -.0085886   -.0013439
      lagemp |   .9785482   .0193408    50.59   0.000     .9403519    1.016744
------------------------------------------------------------------------------
        fips |   absorbed                                     (23 categories)
```

DDD estimates with covariates (GSP), at the state level, in level

```
Regression with robust standard errors                 Number of obs =     355
                                                       F( 22,   303) = 6.1e+05
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9975
                                                       Adj R-squared =  0.9970
                                                       Root MSE      =  20.058
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostddd |  -9.592104   7.062223    -1.36   0.175    -23.48932    4.305108
  laborforce |   6.15e-06   5.63e-06     1.09   0.275    -4.93e-06    .0000172
      laggsp |  -3.48e-06   .0000898    -0.04   0.969    -.0001803    .0001733
        know |  -.0004903   .0020036    -0.24   0.807    -.0044331    .0034525
    lagrdpri |   .0006157   .0018904     0.33   0.745    -.0031042    .0043356
    lagrdfed |   .0011876   .0023793     0.50   0.618    -.0034944    .0058696
      lagemp |   1.004037   .0099336   101.08   0.000      .984489    1.023584
------------------------------------------------------------------------------
        fips |   absorbed                                     (29 categories)
```

DDD estimates with covariates (personal income), at the state level, in level

```
Regression with robust standard errors                 Number of obs =     450
                                                       F( 25,   394) = 2.4e+08
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9967
                                                       Adj R-squared =  0.9963
                                                       Root MSE      =  21.875
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostddd |  -5.825992   7.280549    -0.80   0.424    -20.13957    8.487589
  laborforce |   5.06e-06   3.83e-06     1.32   0.187    -2.48e-06    .0000126
lagperincome |  -6.74e-09   7.43e-08    -0.09   0.928    -1.53e-07    1.39e-07
        know |   .0008211   .0011476     0.72   0.475    -.0014352    .0030773
    lagrdpri |   .0000582   .0007014     0.08   0.934    -.0013208    .0014372
    lagrdfed |   -.000094     .00181    -0.05   0.959    -.0036524    .0034644
      lagemp |    1.00185   .0102791    97.47   0.000     .9816415    1.022059
------------------------------------------------------------------------------
        fips |   absorbed                                     (29 categories)
```

DD estimates with covariates, at the firm level, in level

```
Regression with robust standard errors                 Number of obs =   21006
                                                       F( 31, 20917) = 4013.95
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9914
                                                       Adj R-squared =  0.9913
                                                       Root MSE      =  3.1208
                             (standard errors adjusted for clustering on fips)
```

```
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .1375856   .0711478     1.93   0.053    -.0018697    .2770408
       lagrd |   -.003052   .0028543    -1.07   0.285    -.0086467    .0025427
      lagemp |   1.021873   .0182295    56.06   0.000     .9861422    1.057605
    lagsales |  -.0001528    .000111    -1.38   0.169    -.0003704    .0000647
  knowcounty |    .000018   .0000185     0.97   0.330    -.0000182    .0000542
   emp2digit |   1.29e-08   1.21e-07     0.11   0.915    -2.24e-07    2.50e-07
------------------------------------------------------------------------------
    fips |   absorbed                                        (36 categories)
```

DDD estimates with covariates, at the firm level, in level

```
Regression with robust standard errors                 Number of obs =   35435
                                                       F( 34, 35228) =83383.19
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9855
                                                       Adj R-squared =  0.9854
                                                       Root MSE      =  3.5915
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   tpostddd |  -.0125899   .1508068    -0.08   0.933    -.3081759    .2829961
       lagrd |  -.0010887   .0027771    -0.39   0.695    -.0065318    .0043544
      lagemp |   .9959368   .0204909    48.60   0.000     .9557741      1.0361
    lagsales |  -.0000503   .0001066    -0.47   0.637    -.0002592    .0001586
  knowcounty |  -3.92e-06   .0000104    -0.38   0.707    -.0000244    .0000165
   emp2digit |   7.36e-08   1.14e-07     0.65   0.519    -1.50e-07    2.97e-07
------------------------------------------------------------------------------
    fips |   absorbed                                        (36 categories)
```

DD estimates with covariates (GSP), at the state level, in log

```
Regression with robust standard errors                 Number of obs =     164
                                                       F( 18,   122) = 9312.20
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9914
                                                       Adj R-squared =  0.9885
                                                       Root MSE      =  .13643
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |  -.0363648   .0479478    -0.76   0.450    -.1312821    .0585526
    loglabor |   .7876134   .6031487     1.31   0.194    -.4063798    1.981607
   loglaggsp |  -.3606361   .3583349    -1.01   0.316    -1.069996    .3487237
     logknow |  -.0676621   .1101232    -0.61   0.540    -.2856621    .1503378
 loglagrdpri |   .0532548   .0669012     0.80   0.428    -.0791829    .1856925
 loglagrdfed |  -.0051748   .0068023    -0.76   0.448    -.0186405     .008291
   loglagemp |   .9418852   .0486348    19.37   0.000     .8456077    1.038163
------------------------------------------------------------------------------
    fips |   absorbed                                        (23 categories)
```

DD estimates with covariates (personal income), at the state level, in log

```
Regression with robust standard errors                 Number of obs =     206
                                                       F( 20,   160) = 6240.54
                                                       Prob > F      =  0.0000
```

R-squared = 0.9920
Adj R-squared = 0.9898
Root MSE = .12644

(standard errors adjusted for clustering on fips)

logemp	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
tpostdd	.0017617	.0353166	0.05	0.960	-.0679852	.0715086
loglabor	.7643061	.598686	1.28	0.204	-.4180398	1.946652
loglagperi~e	-.6994304	.6090354	-1.15	0.253	-1.902215	.5033546
logknow	.0665737	.1264108	0.53	0.599	-.1830751	.3162225
loflagrdpri	.0210245	.0490215	0.43	0.669	-.0757881	.1178371
loglagrdfed	-.0067359	.0146418	-0.46	0.646	-.0356521	.0221803
loglagemp	.9399422	.0514448	18.27	0.000	.8383438	1.041541
fips	absorbed					(23 categories)

DDD estimates with covariates (GSP), at the state level, in log

Regression with robust standard errors

Number of obs = 355
F(22, 303) = 7932.67
Prob > F = 0.0000
R-squared = 0.9942
Adj R-squared = 0.9932
Root MSE = .11857

(standard errors adjusted for clustering on fips)

logemp	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
tpostddd	-.0551209	.0501044	-1.10	0.272	-.1537176	.0434758
loglabor	.4129214	.2482973	1.66	0.097	-.075684	.9015268
loglaggsp	-.2591662	.2032932	-1.27	0.203	-.6592115	.140879
logknow	.0385235	.0975973	0.39	0.693	-.1535308	.2305778
loglagrdpri	-.0027681	.0361974	-0.08	0.939	-.0739983	.068462
loglagrdfed	-.0031615	.0052705	-0.60	0.549	-.0135329	.00721
loglagemp	.9713318	.0239539	40.55	0.000	.9241948	1.018469
fips	absorbed					(29 categories)

DDD estimates with covariates (personal income), at the state level, in log

Regression with robust standard errors

Number of obs = 450
F(25, 394) = 78.75
Prob > F = 0.0000
R-squared = 0.9947
Adj R-squared = 0.9939
Root MSE = .11027

(standard errors adjusted for clustering on fips)

logemp	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
tpostddd	-.0249948	.0343479	-0.73	0.467	-.092523	.0425333
loglabor	.3954416	.2143623	1.84	0.066	-.0259953	.8168786
loglagperi~e	-.3859074	.2585302	-1.49	0.136	-.8941785	.1223637
logknow	.0927485	.0746603	1.24	0.215	-.0540339	.2395309
loglagrdpri	-.031224	.0405616	-0.77	0.442	-.1109683	.0485203
loglagrdfed	-.0034993	.0052109	-0.67	0.502	-.0137439	.0067452
loglagemp	.9779401	.0201399	48.56	0.000	.938345	1.017535
fips	absorbed					(29 categories)

DD estimates with covariates, at the firm level, in log

```
Regression with robust standard errors                 Number of obs =   19444
                                                       F( 31, 19355) =31726.75
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9731
                                                       Adj R-squared =  0.9730
                                                       Root MSE      =  .35546
                             (standard errors adjusted for clustering on fips)

------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .0110673   .0142339     0.78   0.437    -.0168325     .038967
 loglagsales |   .0262506   .0055758     4.71   0.000     .0153217    .0371796
logknowcou~y |  -.0019704   .0011905    -1.66   0.098    -.0043038     .000363
   logemp2di |   .0111118   .0047503     2.34   0.019     .0018008    .0204228
   loglagemp |   .9411952   .0070558   133.39   0.000     .9273653    .9550252
    loglagrd |   .0188917   .0041568     4.54   0.000     .0107439    .0270395
------------------------------------------------------------------------------
        fips |   absorbed                                      (36 categories)
```

DDD estimates with covariates, at the firm level, in log

```
Regression with robust standard errors                 Number of obs =   32438
                                                       F( 34, 32240) = 8439.55
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9779
                                                       Adj R-squared =  0.9777
                                                       Root MSE      =  .32953
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   tpostddd |   .0133316   .0172342     0.77   0.439     -.020448    .0471112
 loglagsales |   .0292593   .0044562     6.57   0.000      .020525    .0379936
logknowcou~y |     -.0023   .0011479    -2.00   0.045    -.0045499   -.0000501
   logemp2di |   .0032489   .0035453     0.92   0.359    -.0037001    .0101979
   loglagemp |     .93809   .0056697   165.46   0.000     .9269772    .9492027
    loglagrd |   .0193841   .0025341     7.65   0.000     .0144171    .0243511
------------------------------------------------------------------------------
        fips |   absorbed                                      (36 categories)
```

A.2.4 The matching estimates for the high-technology industry (the firm level)

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .707767738  -.798695749   1.50646349
                        ATT | .707767738  -.003874377   .711642115
----------------------------+-----------------------------------
Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       233 |       233
   Treated |       516 |       516
-----------+-----------+----------
     Total |       749 |       749
```

```
Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .7116421 -.0109238  .3939125  -.0799547   1.503239   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level

```
Probit estimates                                  Number of obs   =        749
                                                  LR chi2(16)     =     500.15
                                                  Prob > chi2     =     0.0000
Log likelihood = -214.27764                       Pseudo R2       =     0.5385
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |   .0290618   .0271093     1.07   0.284    -.0240714    .0821951
       lagrd |   .0002557   .0011371     0.22   0.822     -.001973    .0024844
    lagsales |   -.000141    .000103    -1.37   0.171    -.0003429    .0000608
  knowcounty |    .000299   .0000743     4.02   0.000     .0001534    .0004447
   emp2digit |  -3.25e-07   6.49e-07    -0.50   0.617    -1.60e-06    9.48e-07
 firmsizeall |  -.0158549   .0268756    -0.59   0.555    -.0685301    .0368204
incapita89~n |  -.0000185   .0000608    -0.30   0.761    -.0001376    .0001006
     edu1290 |  -.0060281   .0260051    -0.23   0.817    -.0569972    .0449409
     edu1690 |  -.0798744   .0248816    -3.21   0.001    -.1286415   -.0311073
    sbir8690 |   8.62e-08   6.92e-09    12.46   0.000     7.26e-08    9.97e-08
    univ8690 |  -3.65e-06   3.80e-07    -9.61   0.000    -4.40e-06   -2.91e-06
   popchange |   -1.89604   .4335941    -4.37   0.000    -2.745869   -1.046211
     crime91 |   .0000465   .0000378     1.23   0.219    -.0000275    .0001205
   poverty89 |  -.0944645   .0392973    -2.40   0.016    -.1714858   -.0174431
  unemprat90 |  -.2930074   .0948625    -3.09   0.002    -.4789345   -.1070802
mestdensi~87 |   479.6605   134.7598     3.56   0.000     215.5362    743.7849
       _cons |    4.86232   2.389392     2.03   0.042     .1791975    9.545443
------------------------------------------------------------------------------
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .707767738  -.798695749   1.50646349
                        ATT | 1.00549977   .180948961   .824550808
----------------------------+-----------------------------------
Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        233 |       233
   Treated |       349        167 |       516
-----------+----------------------+----------
     Total |       349        400 |       749

Bootstrap statistics                              Number of obs    =       749
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .8245508  .6534967  1.770796  -2.733997   4.383099   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .707767738  -.798695749   1.50646349
                        ATT | 1.00549977  -.375482044   1.38098181
----------------------------+-----------------------------------
```

```
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        233 |       233
   Treated |       349        167 |       516
-----------+----------------------+----------
     Total |       349        400 |       749

Bootstrap statistics                              Number of obs    =       749
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  1.380982 -.0180164  1.467788  -1.568649   4.330612   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .335860072    .16810281   .167757261
                        ATT | .335860072   .261692933   .074167139
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       233 |       233
   Treated |       516 |       516
-----------+-----------+----------
     Total |       749 |       749
Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0741671 -.0163766  .1395276   -.206224   .3545583   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1, in log

```
Probit estimates                                  Number of obs   =        749
                                                  LR chi2(16)     =     500.15
                                                  Prob > chi2     =     0.0000
Log likelihood = -214.27764                       Pseudo R2       =     0.5385
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |   .0290618   .0271093     1.07   0.284    -.0240714    .0821951
       lagrd |   .0002557   .0011371     0.22   0.822     -.001973    .0024844
    lagsales |   -.000141    .000103    -1.37   0.171    -.0003429    .0000608
  knowcounty |    .000299   .0000743     4.02   0.000     .0001534    .0004447
   emp2digit |  -3.25e-07   6.49e-07    -0.50   0.617    -1.60e-06    9.48e-07
 firmsizeall |  -.0158549   .0268756    -0.59   0.555    -.0685301    .0368204
incapita89~n |  -.0000185   .0000608    -0.30   0.761    -.0001376    .0001006
     edu1290 |  -.0060281   .0260051    -0.23   0.817    -.0569972    .0449409
     edu1690 |  -.0798744   .0248816    -3.21   0.001    -.1286415   -.0311073
    sbir8690 |   8.62e-08   6.92e-09    12.46   0.000     7.26e-08    9.97e-08
    univ8690 |  -3.65e-06   3.80e-07    -9.61   0.000    -4.40e-06   -2.91e-06
   popchange |   -1.89604   .4335941    -4.37   0.000    -2.745869   -1.046211
     crime91 |   .0000465   .0000378     1.23   0.219    -.0000275    .0001205
   poverty89 |  -.0944645   .0392973    -2.40   0.016    -.1714858   -.0174431
  unemprat90 |  -.2930074   .0948625    -3.09   0.002    -.4789345   -.1070802
mestdensi~87 |   479.6605   134.7598     3.56   0.000     215.5362    743.7849
       _cons |    4.86232   2.389392     2.03   0.042     .1791975    9.545443
------------------------------------------------------------------------------
```

```
-------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+--------------------------------------
        logemp_d  Unmatched | .335860072    .16810281   .167757261
                        ATT | .288267927   .244233153   .044034774
----------------------------+--------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        233 |       233
   Treated |       364        152 |       516
-----------+----------------------+----------
     Total |       364        385 |       749
Bootstrap statistics                               Number of obs    =       749
                                                   Replications     =        50
-------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+-----------------------------------------------------------------
       _bs_1 |    50  .0440348  .1690895   .242039  -.4423607   .5304303   (N)
-------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3, in log

```
-------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+--------------------------------------
        logemp_d  Unmatched | .335860072    .16810281   .167757261
                        ATT | .288267927   .163194807    .12507312
----------------------------+--------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        233 |       233
   Treated |       364        152 |       516
-----------+----------------------+----------
     Total |       364        385 |       749
Bootstrap statistics                               Number of obs    =       749
                                                   Replications     =        50
-------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+-----------------------------------------------------------------
       _bs_1 |    50  .1250731   .080807  .1814895  -.2396438     .48979   (N)
-------------------------------------------------------------------------------
```

A.3 THE DETAILED STATISTICAL RESULTS FOR THE ANALYSIS OF R&D SPENDING BY FIRM SIZE

A.3.1 The DD estimates for the all industry

DD estimates with covariates, for small firms, in level

```
Regression with robust standard errors                 Number of obs =   12511
                                                       F( 27, 12295) =41262.11
```

Prob > F = 0.0000
R-squared = 0.7846
Adj R-squared = 0.7809
Root MSE = 1.8165
(standard errors adjusted for clustering on fips)

rd_con	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
tpostdd	.1133616	.0607471	1.87	0.062	-.0057122	.2324355
lagrd	.9821033	.030098	32.63	0.000	.9231064	1.0411
lagsales	.006301	.0015539	4.05	0.000	.003255	.009347
knowcounty	.000031	.0000231	1.34	0.180	-.0000144	.0000763
emp2digit	9.40e-08	9.20e-08	1.02	0.307	-8.64e-08	2.74e-07
unemprte	.0319341	.0143489	2.23	0.026	.003808	.0600601
fips	absorbed					(25 categories)

DD estimates with covariates, for medium firms, in level

Regression with robust standard errors

Number of obs = 5611
F(29, 5416) = 8402.89
Prob > F = 0.0000
R-squared = 0.6097
Adj R-squared = 0.5958
Root MSE = 5.3663
(standard errors adjusted for clustering on fips)

rd_con	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
tpostdd	.9505163	.4156702	2.29	0.022	.1356357	1.765397
lagrd	.6616308	.1618529	4.09	0.000	.344334	.9789276
lagsales	.0321339	.0144477	2.22	0.026	.0038106	.0604572
knowcounty	.0001521	.0000689	2.21	0.027	.0000171	.0002871
emp2digit	-7.65e-08	3.96e-07	-0.19	0.847	-8.53e-07	6.99e-07
unemprte	-.028839	.0589947	-0.49	0.625	-.1444923	.0868143
fips	absorbed					(29 categories)

DD estimates with covariates, for large firms, in level

Regression with robust standard errors

Number of obs = 18670
F(32, 18367) =80123.52
Prob > F = 0.0000
R-squared = 0.9687
Adj R-squared = 0.9682
Root MSE = 38.97
(standard errors adjusted for clustering on fips)

rd_con	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
tpostdd	4.720194	2.034415	2.32	0.020	.7325521	8.707836
lagrd	.9609524	.0347023	27.69	0.000	.8929327	1.028972
lagsales	.0026854	.0013877	1.94	0.053	-.0000346	.0054054
knowcounty	.0002838	.0003296	0.86	0.389	-.0003623	.0009298
emp2digit	3.25e-06	3.23e-06	1.01	0.314	-3.08e-06	9.58e-06
unemprte	-.1750847	.1589872	-1.10	0.271	-.4867144	.136545
fips	absorbed					(28 categories)

DD estimates with covariates, for small firms, in log

```
Regression with robust standard errors                 Number of obs =   10301
                                                       F( 27, 10105) = 5445.11
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8431
                                                       Adj R-squared =  0.8400
                                                       Root MSE      =  .69415
                              (standard errors adjusted for clustering on fips)

------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .0300853   .0358497     0.84   0.401    -.0401873    .1003578
    loglagrd |   .8575081    .009799    87.51   0.000        .8383    .8767161
 loglagsales |   .0294172   .0063008     4.67   0.000     .0170663    .0417681
logknowcou~y |   .0097804   .0024599     3.98   0.000     .0049584    .0146023
   logemp2di |   .0385639   .0182226     2.12   0.034     .0028439    .0742839
    logunemp |   .0858677   .0433095     1.98   0.047     .0009724    .1707629
------------------------------------------------------------------------------
     fips |   absorbed                                      (25 categories)
```

DD estimates with covariates, for medium firms, in log

```
Regression with robust standard errors                 Number of obs =    4782
                                                       F( 65,  4601) =  242.96
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8917
                                                       Adj R-squared =  0.8875
                                                       Root MSE      =  .50792
                         (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |    .089569   .0401377     2.23   0.026     .0108798    .1682582
    loglagrd |   .8758472   .0135652    64.57   0.000     .8492528    .9024415
 loglagsales |   .0017179   .0124748     0.14   0.890    -.0227387    .0261745
logknowcou~y |   .0033768   .0050997     0.66   0.508    -.0066211    .0133747
   logemp2di |  -.0215958   .0186993    -1.15   0.248    -.0582554    .0150637
    logunemp |   -.013766   .0568105    -0.24   0.809    -.1251418    .0976098
------------------------------------------------------------------------------
     fips |   absorbed                                      (29 categories)
```

DD estimates with covariates, for large firms, in log

```
Regression with robust standard errors                 Number of obs =   17989
                                                       F( 32, 17698) =14289.35
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9620
                                                       Adj R-squared =  0.9614
                                                       Root MSE      =   .3797
                              (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .0294979   .0115158     2.56   0.010     .0069258    .0520699
    loglagrd |   .9318294   .0073869   126.15   0.000     .9173503    .9463085
 loglagsales |   .0504978   .0129717     3.89   0.000     .0250719    .0759237
logknowcou~y |   .0030965   .0029508     1.05   0.294    -.0026874    .0088803
   logemp2di |   .0065496   .0053567     1.22   0.221    -.0039499    .0170492
    logunemp |   .0008475   .0186659     0.05   0.964    -.0357395    .0374346
------------------------------------------------------------------------------
     fips |   absorbed                                      (28 categories)
```

A.3.2 The matching estimates for the all industry

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for small firms in the all industry

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | .594163813   .287508432   .306655381
                        ATT | .594163813   .232486352   .361677461
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       131 |       131
   Treated |       252 |       252
-----------+-----------+----------
     Total |       383 |       383

Bootstrap statistics                              Number of obs    =      1411
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .3616775 -.0026897  .1972373  -.0346857   .7580406   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for small firms in the all industry

```
Probit estimates                                  Number of obs   =        383
                                                  LR chi2(15)     =     275.61
                                                  Prob > chi2     =     0.0000
Log likelihood = -108.22373                       Pseudo R2       =     0.5601
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |  -.0042788   .0770625    -0.06   0.956    -.1553186    .1467609
    lagsales |  -.0078537   .0163352    -0.48   0.631    -.0398701    .0241627
  knowcounty |   .0003548   .0001097     3.23   0.001     .0001398    .0005699
   emp2digit |  -1.03e-06   8.48e-07    -1.22   0.224    -2.69e-06    6.30e-07
 firmsizeall |  -.5181844   2.361712    -0.22   0.826    -5.147055    4.110686
incapita89~n |   .0000539   .0000913     0.59   0.555     -.000125    .0002328
     edu1290 |   .0105995    .035197     0.30   0.763    -.0583853    .0795844
     edu1690 |  -.1031326    .036418    -2.83   0.005    -.1745107   -.0317546
    sbir8690 |   9.59e-08   1.01e-08     9.46   0.000     7.60e-08    1.16e-07
    univ8690 |  -4.12e-06   5.44e-07    -7.58   0.000    -5.19e-06   -3.05e-06
   popchange |   -2.08868   .5988593    -3.49   0.000    -3.262423   -.9149375
     crime91 |   .0001614   .0000588     2.75   0.006     .0000462    .0002766
   poverty89 |  -.0988531   .0617542    -1.60   0.109     -.219889    .0221828
  unemprat90 |  -.3895583   .1318051    -2.96   0.003    -.6478915   -.1312251
mestdensi~87 |   758.6466   208.6534     3.64   0.000     349.6934      1167.6
       _cons |   2.545295   3.300163     0.77   0.441    -3.922905    9.013494
------------------------------------------------------------------------------
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | .594163813   .287508432   .306655381
                        ATT |   .5852976   .377328006   .207969594
----------------------------+-----------------------------------
```

```
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        131 |       131
   Treated |       214         38 |       252
-----------+----------------------+----------
     Total |       214        169 |       383

Bootstrap statistics                              Number of obs    =       383
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .2079696  .1809371  .5568832   -.911129   1.327068   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level for small firms in the all industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | .594163813   .287508432   .306655381
                        ATT |   .5852976   .361054089   .224243511
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        131 |       131
   Treated |       214         38 |       252
-----------+----------------------+----------
     Total |       214        169 |       383

Bootstrap statistics                              Number of obs    =       383
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .2242435  .2823874  .5542699  -.8896036   1.338091   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for medium firms in the all industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | 2.46660186   1.24403018   1.22257168
                        ATT | 2.46660186   .697882107   1.76871975
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |        54 |        54
   Treated |       110 |       110
-----------+-----------+----------
     Total |       164 |       164

Bootstrap statistics                              Number of obs    =      1411
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50   1.76872  .0033211   .873484    .013388   3.524052   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for medium firms in the all industry

```
Probit estimates                                  Number of obs   =        164
                                                  LR chi2(15)     =     104.11
                                                  Prob > chi2     =     0.0000
Log likelihood = -51.866591                       Pseudo R2       =     0.5009

------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |  -.0019927   .0342725    -0.06   0.954    -.0691656    .0651801
    lagsales |   -.001523   .0073392    -0.21   0.836    -.0159075    .0128615
  knowcounty |   .0002658   .0001796     1.48   0.139    -.0000862    .0006178
   emp2digit |   1.10e-06   1.77e-06     0.62   0.535    -2.36e-06    4.56e-06
 firmsizeall |  -.2894339      1.802    -0.16   0.872    -3.821289    3.242421
incapita89~n |  -1.02e-06   .0001332    -0.01   0.994    -.0002621      .00026
     edu1290 |  -.0321058   .0565831    -0.57   0.570    -.1430066     .078795
     edu1690 |  -.0254647   .0490342    -0.52   0.604    -.1215699    .0706405
    sbir8690 |   8.49e-08   1.47e-08     5.77   0.000     5.61e-08    1.14e-07
    univ8690 |  -3.98e-06   8.69e-07    -4.58   0.000    -5.68e-06   -2.27e-06
   popchange |  -1.923886   1.217638    -1.58   0.114    -4.310412    .4626398
     crime91 |   .0001027   .0000935     1.10   0.272    -.0000805    .0002858
   poverty89 |  -.1827161   .0925484    -1.97   0.048    -.3641076   -.0013245
  unemprat90 |   .0202572   .2154605     0.09   0.925    -.4020376    .4425521
mestdensi~87 |   62.33725    300.555     0.21   0.836    -526.7397    651.4142
       _cons |   4.996381   5.111004     0.98   0.328    -5.021002    15.01376
------------------------------------------------------------------------------
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
        rd_con_d  Unmatched | 2.46660186   1.24403018   1.22257168
                        ATT | .567138246   3.27737327  -2.71023502
----------------------------+-------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         54 |        54
   Treated |        96         14 |       110
-----------+----------------------+----------
     Total |        96         68 |       164

Bootstrap statistics                              Number of obs    =       164
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -2.710235  4.149492  4.960443  -12.67862   7.258149   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level for medium firms in the all industry

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
        rd_con_d  Unmatched | 2.46660186   1.24403018   1.22257168
                        ATT | .567138246   3.38438312  -2.81724487
----------------------------+-------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         54 |        54
   Treated |        96         14 |       110
-----------+----------------------+----------
     Total |        96         68 |       164
```

```
Bootstrap statistics                              Number of obs    =       164
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -2.817245  3.913313  4.158892  -11.17485   5.540361   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for large firms in the all industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | 33.7646999   38.2249586  -4.46025871
                        ATT | 33.7646999   9.81325249   23.9514474
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       199 |       199
   Treated |       391 |       391
-----------+-----------+----------
     Total |       590 |       590
Bootstrap statistics                              Number of obs    =      1411
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  23.95145  -.413107   6.02501   11.84374   36.05916   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for large firms in the all industry

```
Probit estimates                                  Number of obs   =        590
                                                  LR chi2(15)     =     342.16
                                                  Prob > chi2     =     0.0000
Log likelihood =  -206.0578                       Pseudo R2       =     0.4536
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |   .0012515    .000811     1.54   0.123     -.000338    .0028411
    lagsales |  -.0000873    .000068    -1.28   0.199    -.0002205     .000046
  knowcounty |   .0004081   .0000754     5.41   0.000     .0002604    .0005558
   emp2digit |   3.39e-08   8.82e-07     0.04   0.969    -1.69e-06    1.76e-06
 firmsizeall |  -.0004788   .0054315    -0.09   0.930    -.0111244    .0101669
incapita89~n |  -.0000645   .0000643    -1.00   0.316    -.0001905    .0000616
     edu1290 |   .0295679   .0276154     1.07   0.284    -.0245573    .0836931
     edu1690 |  -.0676436   .0280114    -2.41   0.016    -.1225449   -.0127423
    sbir8690 |   7.95e-08   8.03e-09     9.90   0.000     6.37e-08    9.52e-08
    univ8690 |  -3.10e-06   4.14e-07    -7.49   0.000    -3.91e-06   -2.29e-06
   popchange |  -2.649809   .5760002    -4.60   0.000    -3.778749    -1.52087
     crime91 |   .0001203   .0000363     3.32   0.001     .0000492    .0001914
   poverty89 |    -.11609   .0433863    -2.68   0.007    -.2011255   -.0310544
  unemprat90 |  -.3857884   .1028722    -3.75   0.000    -.5874143   -.1841625
mestdensi~87 |   351.0081   132.2754     2.65   0.008     91.75305    610.2632
       _cons |   2.724281    2.30798     1.18   0.238    -1.799277     7.24784
------------------------------------------------------------------------------
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | 33.7646999   38.2249586  -4.46025871
                        ATT | 40.9764317   10.3975204   30.5789113
----------------------------+-----------------------------------
```

```
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        199 |       199
   Treated |       209        182 |       391
-----------+----------------------+----------
     Total |       209        381 |       590

Bootstrap statistics                              Number of obs    =       590
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  30.57891 -8.910141  18.37183  -6.340666   67.49849   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level for large firms in the all industry

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
        rd_con_d  Unmatched | 33.7646999   38.2249586  -4.46025871
                        ATT | 40.9764317   10.0108069   30.9656249
----------------------------+-------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        199 |       199
   Treated |       209        182 |       391
-----------+----------------------+----------
     Total |       209        381 |       590

Bootstrap statistics                              Number of obs    =       590
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  30.96563 -15.07019  16.62209   -2.43772   64.36897   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log for small firms in the all industry

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
         logrd_d  Unmatched | .185383999  -.020155267   .205539266
                        ATT | .185383999   .043057771   .142326228
----------------------------+-------------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       131 |       131
   Treated |       252 |       252
-----------+-----------+----------
     Total |       383 |       383

Bootstrap statistics                              Number of obs    =      1411
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .1423262  .0206642  .1614128   -.182045   .4666975   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1, in log for small firms in the all industry

```
Probit estimates                                  Number of obs   =        383
                                                  LR chi2(15)     =     275.61
                                                  Prob > chi2     =     0.0000
Log likelihood = -108.22373                       Pseudo R2       =     0.5601

------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |  -.0042788   .0770625    -0.06   0.956    -.1553186    .1467609
    lagsales |  -.0078537   .0163352    -0.48   0.631    -.0398701    .0241627
  knowcounty |   .0003548   .0001097     3.23   0.001     .0001398    .0005699
   emp2digit |  -1.03e-06   8.48e-07    -1.22   0.224    -2.69e-06    6.30e-07
 firmsizeall |  -.5181844   2.361712    -0.22   0.826    -5.147055    4.110686
incapita89~n |   .0000539   .0000913     0.59   0.555     -.000125    .0002328
     edu1290 |   .0105995    .035197     0.30   0.763    -.0583853    .0795844
     edu1690 |  -.1031326    .036418    -2.83   0.005    -.1745107   -.0317546
    sbir8690 |   9.59e-08   1.01e-08     9.46   0.000     7.60e-08    1.16e-07
    univ8690 |  -4.12e-06   5.44e-07    -7.58   0.000    -5.19e-06   -3.05e-06
   popchange |   -2.08868   .5988593    -3.49   0.000    -3.262423   -.9149375
     crime91 |   .0001614   .0000588     2.75   0.006     .0000462    .0002766
   poverty89 |  -.0988531   .0617542    -1.60   0.109     -.219889    .0221828
  unemprat90 |  -.3895583   .1318051    -2.96   0.003    -.6478915   -.1312251
mestdensi~87 |   758.6466   208.6534     3.64   0.000     349.6934      1167.6
       _cons |   2.545295   3.300163     0.77   0.441    -3.922905    9.013494
------------------------------------------------------------------------------
-----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+------------------------------------
         logrd_d  Unmatched | .185383999  -.020155267   .205539266
                        ATT |  .19955527   .281263586  -.081708316
----------------------------+------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        131 |       131
   Treated |       214         38 |       252
-----------+----------------------+----------
     Total |       214        169 |       383

Bootstrap statistics                              Number of obs    =       383
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -.0817083   .339419  .5111875  -1.108978   .9455615   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3, in log for small firms in the all industry

```
-----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+------------------------------------
         logrd_d  Unmatched | .185383999  -.020155267   .205539266
                        ATT |  .19955527   .154167434   .045387836
----------------------------+------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        131 |       131
   Treated |       214         38 |       252
-----------+----------------------+----------
     Total |       214        169 |       383
```

```
Bootstrap statistics                              Number of obs    =        383
                                                  Replications     =         50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0453878   .349979  .4172139  -.7930348   .8838105   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log for medium firms in the all industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
         logrd_d  Unmatched | .431964424   .302355768   .129608656
                        ATT | .431964424   .483041516  -.051077092
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |        54 |        54
   Treated |       110 |       110
-----------+-----------+----------
     Total |       164 |       164
Bootstrap statistics                              Number of obs    =       1411
                                                  Replications     =         50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -.0510771  -.115097  .2391437  -.5316544   .4295002   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1, in log
For medium firms in the all industry

```
Probit estimates                                  Number of obs   =        164
                                                  LR chi2(15)     =     104.11
                                                  Prob > chi2     =     0.0000
Log likelihood = -51.866591                       Pseudo R2       =     0.5009
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |  -.0019927   .0342725    -0.06   0.954    -.0691656    .0651801
    lagsales |   -.001523   .0073392    -0.21   0.836    -.0159075    .0128615
  knowcounty |   .0002658   .0001796     1.48   0.139    -.0000862    .0006178
   emp2digit |   1.10e-06   1.77e-06     0.62   0.535    -2.36e-06    4.56e-06
 firmsizeall |  -.2894339      1.802    -0.16   0.872    -3.821289    3.242421
incapita89~n |  -1.02e-06   .0001332    -0.01   0.994    -.0002621      .00026
     edu1290 |  -.0321058   .0565831    -0.57   0.570    -.1430066     .078795
     edu1690 |  -.0254647   .0490342    -0.52   0.604    -.1215699    .0706405
    sbir8690 |   8.49e-08   1.47e-08     5.77   0.000     5.61e-08    1.14e-07
    univ8690 |  -3.98e-06   8.69e-07    -4.58   0.000    -5.68e-06   -2.27e-06
   popchange |  -1.923886   1.217638    -1.58   0.114    -4.310412    .4626398
     crime91 |   .0001027   .0000935     1.10   0.272    -.0000805    .0002858
   poverty89 |  -.1827161   .0925484    -1.97   0.048    -.3641076   -.0013245
  unemprat90 |   .0202572   .2154605     0.09   0.925    -.4020376    .4425521
mestdensi~87 |   62.33725    300.555     0.21   0.836    -526.7397    651.4142
       _cons |   4.996381   5.111004     0.98   0.328    -5.021002    15.01376
------------------------------------------------------------------------------
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
         logrd_d  Unmatched | .431964424   .302355768   .129608656
                        ATT | .700047945   .489849405    .21019854
----------------------------+-----------------------------------
```

```
Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         54 |        54
   Treated |        96         14 |       110
-----------+----------------------+----------
     Total |        96         68 |       164

Bootstrap statistics                              Number of obs    =       164
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .2101985  -.063182  .5430873  -.8811762   1.301573   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3, in log for medium firms in the all industry

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
         logrd_d  Unmatched | .431964424   .302355768   .129608656
                        ATT | .700047945   .616010731   .084037214
----------------------------+-------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         54 |        54
   Treated |        96         14 |       110
-----------+----------------------+----------
     Total |        96         68 |       164

Bootstrap statistics                              Number of obs    =       164
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0840372  .0084981  .4305536  -.7811927   .9492671   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log for large firms in the all industry

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
         logrd_d  Unmatched | .726580633   .424573475   .302007158
                        ATT | .726580633   .473503296   .253077337
----------------------------+-------------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       199 |       199
   Treated |       391 |       391
-----------+-----------+----------
     Total |       590 |       590

Bootstrap statistics                              Number of obs    =      1411
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .2530773 -.0448756   .199474  -.1477807   .6539354   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1, in log for large firms in the all industry

Probit estimates — Number of obs = 590; LR chi2(15) = 342.16; Prob > chi2 = 0.0000; Pseudo R2 = 0.4536

Log likelihood = -206.0578

tst	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lagrd	.0012515	.000811	1.54	0.123	-.000338	.0028411
lagsales	-.0000873	.000068	-1.28	0.199	-.0002205	.000046
knowcounty	.0004081	.0000754	5.41	0.000	.0002604	.0005558
emp2digit	3.39e-08	8.82e-07	0.04	0.969	-1.69e-06	1.76e-06
firmsizeall	-.0004788	.0054315	-0.09	0.930	-.0111244	.0101669
incapita89~n	-.0000645	.0000643	-1.00	0.316	-.0001905	.0000616
edu1290	.0295679	.0276154	1.07	0.284	-.0245573	.0836931
edu1690	-.0676436	.0280114	-2.41	0.016	-.1225449	-.0127423
sbir8690	7.95e-08	8.03e-09	9.90	0.000	6.37e-08	9.52e-08
univ8690	-3.10e-06	4.14e-07	-7.49	0.000	-3.91e-06	-2.29e-06
popchange	-2.649809	.5760002	-4.60	0.000	-3.778749	-1.52087
crime91	.0001203	.0000363	3.32	0.001	.0000492	.0001914
poverty89	-.11609	.0433863	-2.68	0.007	-.2011255	-.0310544
unemprat90	-.3857884	.1028722	-3.75	0.000	-.5874143	-.1841625
mestdensi~87	351.0081	132.2754	2.65	0.008	91.75305	610.2632
_cons	2.724281	2.30798	1.18	0.238	-1.799277	7.24784

Variable	Sample	Treated	Controls	Difference
logrd_d	Unmatched	.726580633	.424573475	.302007158
	ATT	.491202765	.454235166	.036967599

Treatment assignment	support Off suppo	On suppor	Total
Untreated	0	199	199
Treated	273	118	391
Total	273	317	590

Bootstrap statistics — Number of obs = 590; Replications = 50

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf.	Interval]	
_bs_1	50	.0369676	.0761927	.2445866	-.4545476	.5284828	(N)

Matching estimates with propensity score metrics, with caliper, # of match=3, in log for large firms in the all industry

Variable	Sample	Treated	Controls	Difference
logrd_d	Unmatched	.726580633	.424573475	.302007158
	ATT	.491202765	.406762475	.08444029

Treatment assignment	support Off suppo	On suppor	Total
Untreated	0	199	199
Treated	273	118	391
Total	273	317	590

```
Bootstrap statistics                               Number of obs    =       590
                                                   Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0844403  .0095726  .2323215  -.3824273   .5513078   (N)
------------------------------------------------------------------------------
```

A.3.3 The DD/DDD estimates for the high-technology industry

DD estimates with covariates, for small firms, in level

```
Regression with robust standard errors                      Number of obs =    9759
                                                            F( 26,  9646) = 1278.23
                                                            Prob > F      =  0.0000
                                                            R-squared     =  0.7814
                                                            Adj R-squared =  0.7788
                                                            Root MSE      =  2.0169
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      rd_con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .1251549    .072948     1.72   0.086    -.0178384    .2681482
       lagrd |   .9852522   .0304475    32.36   0.000     .9255686    1.044936
    lagsales |   .0066757   .0017596     3.79   0.000     .0032266    .0101248
  knowcounty |   .0000363   .0000252     1.44   0.149     -.000013    .0000857
   emp2digit |   9.07e-08   9.98e-08     0.91   0.363    -1.05e-07    2.86e-07
    unemprte |   .0360196   .0170289     2.12   0.034     .0026394    .0693998
------------------------------------------------------------------------------
        fips |   absorbed                                     (25 categories)
```

DD estimates with covariates, for medium firms, in level

```
Regression with robust standard errors                      Number of obs =    3760
                                                            F( 28,  3651) = 1391.63
                                                            Prob > F      =  0.0000
                                                            R-squared     =  0.5978
                                                            Adj R-squared =  0.5859
                                                            Root MSE      =  6.4471
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      rd_con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   1.298814   .5840186     2.22   0.026     .1537794     2.44385
       lagrd |   .6362317   .1624404     3.92   0.000     .3177487    .9547147
    lagsales |    .044771   .0183513     2.44   0.015     .0087912    .0807507
  knowcounty |   .0002118   .0000759     2.79   0.005      .000063    .0003607
   emp2digit |   1.22e-07   4.61e-07     0.26   0.792    -7.81e-07    1.02e-06
    unemprte |  -.0473858   .0934491    -0.51   0.612    -.2306033    .1358318
------------------------------------------------------------------------------
        fips |   absorbed                                     (29 categories)
```

DD estimates with covariates, for large firms, in level

```
Regression with robust standard errors                      Number of obs =    8708
                                                            F( 27,  8585) = 2168.31
                                                            Prob > F      =  0.0000
                                                            R-squared     =  0.9689
                                                            Adj R-squared =  0.9685
                                                            Root MSE      =  54.344
```

```
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      rd_con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   8.891604   4.356379     2.04   0.041     .3520534    17.43115
       lagrd |   .9512156   .0398624    23.86   0.000     .8730756    1.029356
    lagsales |   .0033548   .0016223     2.07   0.039     .0001747     .006535
  knowcounty |   .0007426   .0005313     1.40   0.162    -.0002988     .001784
   emp2digit |   2.10e-06   4.80e-06     0.44   0.663    -7.32e-06    .0000115
    unemprte |  -.2423755   .3078382    -0.79   0.431    -.8458124    .3610613
------------------------------------------------------------------------------
        fips |   absorbed                                     (28 categories)
```

DDD estimates with covariates, for small firms, in level

```
Regression with robust standard errors                Number of obs =   12511
                                                      F( 27, 12292) =12118.29
                                                      Prob > F      =  0.0000
                                                      R-squared     =  0.7847
                                                      Adj R-squared =  0.7809
                                                      Root MSE      =  1.8163
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      rd_con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   tpostddd |   .1182074   .0880104     1.34   0.179    -.0543069    .2907217
       lagrd |   .9809829   .0303008    32.37   0.000     .9215885    1.040377
    lagsales |   .0062676   .0015498     4.04   0.000     .0032299    .0093054
  knowcounty |     .00003   .0000229     1.31   0.190    -.0000148    .0000748
   emp2digit |   8.49e-08   8.86e-08     0.96   0.338    -8.88e-08    2.59e-07
    unemprte |   .0318074   .0144179     2.21   0.027      .003546    .0600688
------------------------------------------------------------------------------
        fips |   absorbed                                     (25 categories)
```

DDD estimates with covariates, for medium firms, in level

```
Regression with robust standard errors                Number of obs =    5611
                                                      F( 29,  5413) =  649.60
                                                      Prob > F      =  0.0000
                                                      R-squared     =  0.6101
                                                      Adj R-squared =  0.5959
                                                      Root MSE      =  5.3657
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      rd_con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   tpostddd |   .7734504   .4129345     1.87   0.061    -.0360674    1.582968
       lagrd |   .6609191   .1619072     4.08   0.000     .3435158    .9783223
    lagsales |   .0321108   .0144897     2.22   0.027     .0037052    .0605164
  knowcounty |   .0001465   .0000674     2.17   0.030     .0000143    .0002787
   emp2digit |  -1.06e-07   3.96e-07    -0.27   0.788    -8.83e-07    6.71e-07
    unemprte |  -.0326172   .0588707    -0.55   0.580    -.1480276    .0827931
------------------------------------------------------------------------------
        fips |   absorbed                                     (29 categories)
```

DDD estimates with covariates, for large firms, in level

```
Regression with robust standard errors                Number of obs =   18670
                                                      F( 32, 18364) =31414.83
                                                      Prob > F      =  0.0000
                                                      R-squared     =  0.9688
                                                      Adj R-squared =  0.9682
                                                      Root MSE      =  38.959
                             (standard errors adjusted for clustering on fips)
```

```
------------------------------------------------------------------------------
             |               Robust
      rd_con |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostddd |   7.193452    3.83939     1.87   0.061    -.3321098    14.71901
       lagrd |   .9604343   .0346083    27.75   0.000     .8925988     1.02827
    lagsales |   .0027131   .0013791     1.97   0.049     9.97e-06    .0054162
  knowcounty |   .0002497   .0003031     0.82   0.410    -.0003444    .0008438
   emp2digit |   2.89e-06   3.31e-06     0.87   0.382    -3.60e-06    9.39e-06
    unemprte |  -.2167695   .1575355    -1.38   0.169    -.5255539    .0920148
------------------------------------------------------------------------------
        fips |   absorbed                                      (28 categories)
```

DD estimates with covariates, for small firms, in log

```
Regression with robust standard errors                 Number of obs =    8233
                                                       F( 26,  8121) =32107.77
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8478
                                                       Adj R-squared =  0.8457
                                                       Root MSE      =  .65934
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |  -.0198726   .0398493    -0.50   0.618    -.0979873    .0582422
    loglagrd |   .8670516   .0128395    67.53   0.000      .841883    .8922203
 loglagsales |   .0238669   .0069566     3.43   0.001     .0102302    .0375037
logknowcou~y |   .0110651    .004224     2.62   0.009     .0027849    .0193452
   logemp2di |   .0339025    .017812     1.90   0.057    -.0010135    .0688186
     logunemp |   .0977914   .0495362     1.97   0.048     .0006878     .194895
------------------------------------------------------------------------------
        fips |   absorbed                                      (25 categories)
```

DD estimates with covariates, for medium firms, in log

```
Regression with robust standard errors                 Number of obs =    3286
                                                       F( 36,  3180) = 2422.90
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8822
                                                       Adj R-squared =  0.8783
                                                       Root MSE      =  .48238
                              (standard errors adjusted for clustering on sic)
------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   .1391311   .0503601     2.76   0.006     .0403895    .2378728
    loglagrd |   .8790151   .0216014    40.69   0.000      .836661    .9213692
 loglagsales |  -.0124507   .0276478    -0.45   0.653    -.0666601    .0417586
logknowcou~y |   .0011591   .0044504     0.26   0.795    -.0075668    .0098849
   logemp2di |  -.0299767   .0221939    -1.35   0.177    -.0734925    .0135391
    logunemp |  -.0466552   .0610745    -0.76   0.445    -.1664047    .0730943
------------------------------------------------------------------------------
        fips |   absorbed                                      (29 categories)
```

DD estimates with covariates, for large firms, in log

```
Regression with robust standard errors                 Number of obs =    8631
                                                       F( 27,  8508) =  811.34
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9652
                                                       Adj R-squared =  0.9647
                                                       Root MSE      =  .36397
                             (standard errors adjusted for clustering on fips)
```

```
------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .0099212   .0182797     0.54   0.587    -.0259115    .0457539
    loglagrd |   .9455339   .0070233   134.63   0.000     .9317665    .9593013
 loglagsales |   .0240297   .0127735     1.88   0.060    -.0010094    .0490689
logknowcou~y |   .0089178   .0045805     1.95   0.052    -.0000611    .0178967
   logemp2di |   .0135948   .0105027     1.29   0.196    -.0069931    .0341827
    logunemp |   .0039333   .0214418     0.18   0.854    -.0380979    .0459644
------------------------------------------------------------------------------
        fips |   absorbed                                     (28 categories)
```

DDD estimates with covariates, for small firms, in log

```
Regression with robust standard errors                 Number of obs =   10301
                                                       F( 27, 10102) =14964.87
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8434
                                                       Adj R-squared =  0.8403
                                                       Root MSE      =  .69356
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   tpostddd |  -.2057073   .1123555    -1.83   0.067    -.4259464    .0145318
    loglagrd |   .8565214   .0098086    87.32   0.000     .8372947    .8757482
 loglagsales |    .029491   .0062871     4.69   0.000     .0171671     .041815
logknowcou~y |   .0095516    .002413     3.96   0.000     .0048217    .0142815
   logemp2di |   .0387386   .0178451     2.17   0.030     .0037587    .0737185
    logunemp |    .089898   .0436484     2.06   0.039     .0043386    .1754575
------------------------------------------------------------------------------
        fips |   absorbed                                     (25 categories)
```

DDD estimates with covariates, for medium firms, in log

```
Regression with robust standard errors                 Number of obs =    5148
                                                       F( 28,  4965) = 8724.42
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8853
                                                       Adj R-squared =  0.8811
                                                       Root MSE      =  .52704
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   tpostddd |   .0581522   .0830679     0.70   0.484    -.1046975    .2210019
    loglagrd |    .868149    .013442    64.58   0.000     .8417967    .8945013
 loglagsales |  -.0120449   .0116583    -1.03   0.302    -.0349005    .0108106
logknowcou~y |   .0021274   .0076405     0.28   0.781    -.0128513     .017106
   logemp2di |  -.0000997   .0183782    -0.01   0.996     -.036129    .0359296
    logunemp |  -.0644002   .0434937    -1.48   0.139    -.1496671    .0208667
------------------------------------------------------------------------------
        fips |   absorbed                                     (29 categories)
```

DDD estimates with covariates, for large firms, in log

```
Regression with robust standard errors                 Number of obs =   17989
                                                       F( 32, 17695) =13417.45
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9620
                                                       Adj R-squared =  0.9614
                                                       Root MSE      =  .37971
                             (standard errors adjusted for clustering on fips)
```

```
------------------------------------------------------------------------------
             |               Robust
       logrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostddd |  -.0181133   .0288002    -0.63   0.529    -.0745645    .0383379
    loglagrd |   .9319544   .0073498   126.80   0.000     .9175481    .9463608
 loglagsales |   .0504972   .0128759     3.92   0.000     .0252591    .0757353
logknowcou~y |   .0031679   .0030265     1.05   0.295    -.0027644    .0091001
    logemp2di |    .006591   .0052323     1.26   0.208    -.0036647    .0168468
    logunemp |   .0024004   .0184071     0.13   0.896    -.0336793    .0384801
------------------------------------------------------------------------------
        fips |   absorbed                                      (28 categories)
```

A.3.4 The matching estimates for the high-technology industry

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for small firms in the high-technology industry

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | .758375458   .366394785   .391980673
                        ATT | .758375458   .347413602   .410961856
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       101 |       101
   Treated |       202 |       202
-----------+-----------+----------
     Total |       303 |       303
Bootstrap statistics                              Number of obs    =      1411
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .4109619   .081072  .2084909  -.0080162   .8299399   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for small firms in the high-technology industry

```
Probit estimates                                  Number of obs   =        303
                                                  LR chi2(15)     =     216.78
                                                  Prob > chi2     =     0.0000
Log likelihood = -84.471363                       Pseudo R2       =     0.5620
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |  -.0253597   .0781978    -0.32   0.746    -.1786246    .1279052
    lagsales |   .0087409   .0167527     0.52   0.602    -.0240937    .0415755
  knowcounty |   .0001894   .0001304     1.45   0.146    -.0000661    .0004449
   emp2digit |  -1.40e-06   9.92e-07    -1.41   0.159    -3.34e-06    5.48e-07
 firmsizeall |  -1.892086   2.642807    -0.72   0.474    -7.071893    3.287722
incapita89~n |   .0000495   .0001021     0.48   0.628    -.0001506    .0002496
     edu1290 |  -.0451521   .0442782    -1.02   0.308    -.1319358    .0416315
     edu1690 |  -.0676485   .0398811    -1.70   0.090     -.145814     .010517
    sbir8690 |   9.55e-08   1.18e-08     8.08   0.000     7.23e-08    1.19e-07
    univ8690 |  -3.97e-06   6.16e-07    -6.45   0.000    -5.18e-06   -2.77e-06
   popchange |  -1.511105   .6493031    -2.33   0.020    -2.783716   -.2384943
     crime91 |   .0002455   .0000741     3.31   0.001     .0001003    .0003907
```

```
   poverty89 |   -.145368   .0717313    -2.03   0.043    -.2859589   -.0047772
  unemprat90 |  -.3790756   .1625847    -2.33   0.020    -.6977357   -.0604156
mestdensi~87 |   1096.207   274.7441     3.99   0.000      557.718    1634.695
       _cons |   5.411083   4.106536     1.32   0.188    -2.637579    13.45975
------------------------------------------------------------------------------
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | .758375458   .366394785   .391980673
                        ATT | 1.00004478   .426808845   .573235936
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        101 |       101
   Treated |       148         54 |       202
-----------+----------------------+----------
     Total |       148        155 |       303
Bootstrap statistics                              Number of obs    =       303
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .5732359  .1862734  .6071916  -.6469612   1.793433   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level for small firms in the high-technology industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | .758375458   .366394785   .391980673
                        ATT | 1.00004478   .217849844   .782194938
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        101 |       101
   Treated |       148         54 |       202
-----------+----------------------+----------
     Total |       148        155 |       303
Bootstrap statistics                              Number of obs    =       303
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .7821949   .113054  .6282165  -.4802533   2.044643   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for medium firms in the high-technology industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | 2.96977869   2.13587907   .833899622
                        ATT | 2.96977869   .603674835   2.36610385
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |        30 |        30
   Treated |        84 |        84
-----------+-----------+----------
     Total |       114 |       114
```

```
Bootstrap statistics                              Number of obs    =      1411
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  2.366104 -.2473935  1.492487  -.6331619    5.36537   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for medium firms in the high-technology industry

```
Probit estimates                                  Number of obs   =        114
                                                  LR chi2(15)     =      74.78
                                                  Prob > chi2     =     0.0000
Log likelihood = -28.311935                       Pseudo R2       =     0.5691
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |  -.0044462   .0410222    -0.11   0.914    -.0848482    .0759557
    lagsales |  -.0000915    .011357    -0.01   0.994    -.0223509    .0221679
  knowcounty |   .0005643   .0002733     2.06   0.039     .0000287    .0010999
   emp2digit |   1.60e-06   2.40e-06     0.66   0.506    -3.11e-06    6.30e-06
 firmsizeall |  -1.917604   2.913354    -0.66   0.510    -7.627674    3.792466
incapita89~n |  -3.37e-06   .0001756    -0.02   0.985    -.0003476    .0003408
     edu1290 |   .0130812   .0765736     0.17   0.864    -.1370002    .1631626
     edu1690 |  -.1007147   .0734226    -1.37   0.170    -.2446204     .043191
    sbir8690 |   1.10e-07   2.49e-08     4.40   0.000     6.07e-08    1.59e-07
    univ8690 |  -5.34e-06   1.42e-06    -3.77   0.000    -8.12e-06   -2.56e-06
   popchange |  -1.823142   1.273054    -1.43   0.152    -4.318282    .6719976
     crime91 |   .0000177   .0001293     0.14   0.891    -.0002357     .000271
   poverty89 |  -.1493721    .114046    -1.31   0.190    -.3728982     .074154
  unemprat90 |  -.3688869   .3640206    -1.01   0.311    -1.082354    .3445803
mestdensi~87 |  -13.88161   428.3147    -0.03   0.974    -853.3629    825.5997
       _cons |   6.383084   7.502602     0.85   0.395    -8.321746    21.08791
------------------------------------------------------------------------------
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | 2.96977869   2.13587907   .833899622
                        ATT | 4.38304795   .530225536   3.85282241
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         30 |        30
   Treated |        79          5 |        84
-----------+----------------------+----------
     Total |        79         35 |       114

Bootstrap statistics                              Number of obs    =       114
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    32  3.852822 -.0944126  8.484887  -13.45222   21.15786   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level for medium firms in the high-technology industry

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | 2.96977869   2.13587907   .833899622
                        ATT | 4.38304795   .533855417   3.84919253
----------------------------+-----------------------------------
```

```
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         30 |        30
   Treated |        79          5 |        84
-----------+----------------------+----------
     Total |        79         35 |       114

Bootstrap statistics                              Number of obs    =       114
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    30  3.849193 -2.495623  4.803407  -5.974877   13.67326   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for large firms in the high-technology industry

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | 51.6139034   85.3704754  -33.7565721
                        ATT | 51.6139034   15.1816737   36.4322297
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |        85 |        85
   Treated |       223 |       223
-----------+-----------+----------
     Total |       308 |       308

Bootstrap statistics                              Number of obs    =      1411
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  36.43223 -1.685795  10.64171   15.04691   57.81754   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for large firms in the high-technology industry

```
Probit estimates                                  Number of obs   =        308
                                                  LR chi2(15)     =     204.14
                                                  Prob > chi2     =     0.0000
Log likelihood = -79.375542                       Pseudo R2       =     0.5625
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |   .0011582   .0011494     1.01   0.314    -.0010946    .0034109
    lagsales |  -.0001078   .0000946    -1.14   0.255    -.0002933    .0000776
  knowcounty |   .0003219    .000118     2.73   0.006     .0000906    .0005532
   emp2digit |  -9.30e-08   1.29e-06    -0.07   0.943    -2.62e-06    2.44e-06
 firmsizeall |    .003224   .0134687     0.24   0.811    -.0231742    .0296222
incapita89~n |  -.0000826   .0001022    -0.81   0.419    -.0002829    .0001178
     edu1290 |    .026488    .042807     0.62   0.536    -.0574121    .1103882
     edu1690 |  -.0783814   .0414033    -1.89   0.058    -.1595304    .0027676
    sbir8690 |   8.55e-08   1.22e-08     7.00   0.000     6.15e-08    1.09e-07
    univ8690 |  -3.43e-06   6.37e-07    -5.39   0.000    -4.68e-06   -2.18e-06
   popchange |  -2.174384   .7638144    -2.85   0.004    -3.671433   -.6773349
     crime91 |  -.0000536   .0000607    -0.88   0.377    -.0001725    .0000653
   poverty89 |  -.0894627   .0682773    -1.31   0.190    -.2232837    .0443584
  unemprat90 |  -.2591847    .165915    -1.56   0.118    -.5843721    .0660026
```

```
mestdensi~87 |    345.393   215.2018     1.60   0.108    -76.39468    767.1807
       _cons |   3.774186   3.960143     0.95   0.341    -3.987551    11.53592
------------------------------------------------------------------------------
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | 51.6139034   85.3704754  -33.7565721
                        ATT | 42.5114513   26.8823001   15.6291512
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         85 |        85
   Treated |       190         33 |       223
-----------+----------------------+----------
     Total |       190        118 |       308

Bootstrap statistics                               Number of obs    =       308
                                                   Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  15.62915  24.40521  62.98355   -110.941   142.1993   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level for large firms in the high-technology industry

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        rd_con_d  Unmatched | 51.6139034   85.3704754  -33.7565721
                        ATT | 42.5114513   26.9267195   15.5847317
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         85 |        85
   Treated |       190         33 |       223
-----------+----------------------+----------
     Total |       190        118 |       308

Bootstrap statistics                               Number of obs    =       308
                                                   Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  15.58473   7.98411  45.60711  -76.06618   107.2356   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log for small firms in the high-technology industry

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
         logrd_d  Unmatched | .300138511   .097517101   .202621409
                        ATT | .300138511   .103291052   .196847458
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       101 |       101
   Treated |       202 |       202
-----------+-----------+----------
     Total |       303 |       303
```

```
Bootstrap statistics                               Number of obs    =      1411
                                                   Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .1968475  .0392637  .2211308  -.2475316   .6412265   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1, in log for small firms in the high-technology industry

```
Probit estimates                                   Number of obs   =        303
                                                   LR chi2(15)     =     216.78
                                                   Prob > chi2     =     0.0000
Log likelihood = -84.471363                        Pseudo R2       =     0.5620
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lagrd |  -.0253597   .0781978    -0.32   0.746    -.1786246    .1279052
    lagsales |   .0087409   .0167527     0.52   0.602    -.0240937    .0415755
  knowcounty |   .0001894   .0001304     1.45   0.146    -.0000661    .0004449
   emp2digit |  -1.40e-06   9.92e-07    -1.41   0.159    -3.34e-06    5.48e-07
 firmsizeall |  -1.892086   2.642807    -0.72   0.474    -7.071893    3.287722
incapita89~n |   .0000495   .0001021     0.48   0.628    -.0001506    .0002496
     edu1290 |  -.0451521   .0442782    -1.02   0.308    -.1319358    .0416315
     edu1690 |  -.0676485   .0398811    -1.70   0.090     -.145814     .010517
    sbir8690 |   9.55e-08   1.18e-08     8.08   0.000     7.23e-08    1.19e-07
    univ8690 |  -3.97e-06   6.16e-07    -6.45   0.000    -5.18e-06   -2.77e-06
   popchange |  -1.511105   .6493031    -2.33   0.020    -2.783716   -.2384943
     crime91 |   .0002455   .0000741     3.31   0.001     .0001003    .0003907
   poverty89 |   -.145368   .0717313    -2.03   0.043    -.2859589   -.0047772
  unemprat90 |  -.3790756   .1625847    -2.33   0.020    -.6977357   -.0604156
mestdensi~87 |   1096.207   274.7441     3.99   0.000      557.718    1634.695
       _cons |   5.411083   4.106536     1.32   0.188    -2.637579    13.45975
------------------------------------------------------------------------------
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
         logrd_d  Unmatched | .300138511   .097517101   .202621409
                        ATT | .435281152   .410701829   .024579322
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        101 |       101
   Treated |       173         29 |       202
-----------+----------------------+----------
     Total |       173        130 |       303

Bootstrap statistics                               Number of obs    =       303
                                                   Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0245793  .5743329  .4132917  -.8059614     .85512   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3, in log for small firms in the high-technology industry

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
         logrd_d  Unmatched | .300138511   .097517101   .202621409
                        ATT | .435281152   .290910948   .144370203
----------------------------+-----------------------------------
```

Treatment assignment	support: Off suppo	On suppor	Total
Untreated	0	101	101
Treated	173	29	202
Total	173	130	303

Bootstrap statistics — Number of obs = 303; Replications = 50

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf. Interval]		
_bs_1	50	.1443702	.2549973	.6047759	-1.070972	1.359713	(N)

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log for medium firms in the high-technology industry

Variable	Sample	Treated	Controls	Difference
logrd_d	Unmatched	.463940505	.366281539	.097658966
	ATT	.463940505	.189505009	.274435496

Treatment assignment	support: On suppor	Total
Untreated	30	30
Treated	84	84
Total	114	114

Bootstrap statistics — Number of obs = 1411; Replications = 50

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf. Interval]		
_bs_1	50	.2744355	.0363102	.416849	-.563254	1.112125	(N)

Matching estimates with propensity score metrics, with caliper, # of match=1, in log for medium firms in the high-technology industry

Probit estimates — Number of obs = 114; LR chi2(15) = 74.78; Prob > chi2 = 0.0000

Log likelihood = -28.311935 — Pseudo R2 = 0.5691

tst	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
lagrd	-.0044462	.0410222	-0.11	0.914	-.0848482	.0759557
lagsales	-.0000915	.011357	-0.01	0.994	-.0223509	.0221679
knowcounty	.0005643	.0002733	2.06	0.039	.0000287	.0010999
emp2digit	1.60e-06	2.40e-06	0.66	0.506	-3.11e-06	6.30e-06
firmsizeall	-1.917604	2.913354	-0.66	0.510	-7.627674	3.792466
incapita89~n	-3.37e-06	.0001756	-0.02	0.985	-.0003476	.0003408
edu1290	.0130812	.0765736	0.17	0.864	-.1370002	.1631626
edu1690	-.1007147	.0734226	-1.37	0.170	-.2446204	.043191
sbir8690	1.10e-07	2.49e-08	4.40	0.000	6.07e-08	1.59e-07
univ8690	-5.34e-06	1.42e-06	-3.77	0.000	-8.12e-06	-2.56e-06
popchange	-1.823142	1.273054	-1.43	0.152	-4.318282	.6719976
crime91	.0000177	.0001293	0.14	0.891	-.0002357	.000271
poverty89	-.1493721	.114046	-1.31	0.190	-.3728982	.074154
unemprat90	-.3688869	.3640206	-1.01	0.311	-1.082354	.3445803
mestdensi~87	-13.88161	428.3147	-0.03	0.974	-853.3629	825.5997
_cons	6.383084	7.502602	0.85	0.395	-8.321746	21.08791

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
         logrd_d  Unmatched | .463940505   .366281539   .097658966
                        ATT | .659553366   .483248484   .176304882
----------------------------+-------------------------------------
 Treatment |         support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         30 |        30
   Treated |        74         10 |        84
-----------+----------------------+----------
     Total |        74         40 |       114

Bootstrap statistics                              Number of obs    =       114
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    39  .1763049  .3182976  .6579183  -1.155581   1.508191   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3, in log for medium firms in the high-technology industry

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
         logrd_d  Unmatched | .463940505   .366281539   .097658966
                        ATT | .659553366   .553747055   .105806311
----------------------------+-------------------------------------
 Treatment |         support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         30 |        30
   Treated |        74         10 |        84
-----------+----------------------+----------
     Total |        74         40 |       114

Bootstrap statistics                              Number of obs    =       114
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    37  .1058063  .1727817  .6884994  -1.290535   1.502148   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log for large firms in the high-technology industry

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
         logrd_d  Unmatched |  .95369657   .759208126   .194488444
                        ATT |  .95369657   1.02633296  -.072636393
----------------------------+-------------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |        85 |        85
   Treated |       223 |       223
-----------+-----------+----------
     Total |       308 |       308
```

Bootstrap statistics — Number of obs = 1411; Replications = 50

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf.	Interval]	
_bs_1	50	-.0726364	.0899906	.1729332	-.4201587	.2748859	(N)

Matching estimates with propensity score metrics, with caliper, # of match=1, in log for large firms in the high-technology industry

Probit estimates — Number of obs = 308; LR chi2(15) = 204.14; Prob > chi2 = 0.0000

Log likelihood = -79.375542 — Pseudo R2 = 0.5625

tst	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lagrd	.0011582	.0011494	1.01	0.314	-.0010946	.0034109
lagsales	-.0001078	.0000946	-1.14	0.255	-.0002933	.0000776
knowcounty	.0003219	.000118	2.73	0.006	.0000906	.0005532
emp2digit	-9.30e-08	1.29e-06	-0.07	0.943	-2.62e-06	2.44e-06
firmsizeall	.003224	.0134687	0.24	0.811	-.0231742	.0296222
incapita89~n	-.0000826	.0001022	-0.81	0.419	-.0002829	.0001178
edu1290	.026488	.042807	0.62	0.536	-.0574121	.1103882
edu1690	-.0783814	.0414033	-1.89	0.058	-.1595304	.0027676
sbir8690	8.55e-08	1.22e-08	7.00	0.000	6.15e-08	1.09e-07
univ8690	-3.43e-06	6.37e-07	-5.39	0.000	-4.68e-06	-2.18e-06
popchange	-2.174384	.7638144	-2.85	0.004	-3.671433	-.6773349
crime91	-.0000536	.0000607	-0.88	0.377	-.0001725	.0000653
poverty89	-.0894627	.0682773	-1.31	0.190	-.2232837	.0443584
unemprat90	-.2591847	.165915	-1.56	0.118	-.5843721	.0660026
mestdensi~87	345.393	215.2018	1.60	0.108	-76.39468	767.1807
_cons	3.774186	3.960143	0.95	0.341	-3.987551	11.53592

Variable	Sample	Treated	Controls	Difference
logrd_d	Unmatched	.95369657	.759208126	.194488444
	ATT	.766964945	.649216056	.117748889

Treatment assignment	support: Off suppo	On suppor	Total
Untreated	0	85	85
Treated	190	33	223
Total	190	118	308

Bootstrap statistics — Number of obs = 308; Replications = 50

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf.	Interval]	
_bs_1	50	.1177489	-.0042366	.4667617	-.8202439	1.055742	(N)

Matching estimates with propensity score metrics, with caliper, # of match=3, in log for large firms in the high-technology industry

Variable	Sample	Treated	Controls	Difference
logrd_d	Unmatched	.95369657	.759208126	.194488444
	ATT	.766964945	.651241806	.115723139

Treatment assignment	support: Off suppo	On suppor	Total
Untreated	0	85	85
Treated	190	33	223
Total	190	118	308

```
Bootstrap statistics                              Number of obs    =       308
                                                  Replications     =        50

------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .1157231 -.0477548  .4581183  -.8049001   1.036346   (N)
------------------------------------------------------------------------------
```

A.4 THE DETAILED STATISTICAL ESTIMATES FOR THE ANALYSIS OF EMPLOYMENT BY FIRM SIZE

A.4.1 The DD estimates for the all industry

DD estimates with covariates, for small firms, in level

```
Regression with robust standard errors                 Number of obs =   11566
                                                       F( 26, 11353) =27932.34
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7539
                                                       Adj R-squared =  0.7493
                                                       Root MSE      =  .03843
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .0011416     .00286     0.40   0.690    -.0044646    .0067477
       lagrd |   .0005139   .0003141     1.64   0.102    -.0001017    .0011296
      lagemp |   .8349938   .0296015    28.21   0.000     .7769697    .8930179
    lagsales |   .0005004   .0000701     7.14   0.000      .000363    .0006379
  knowcounty |  -1.84e-07   2.29e-07    -0.81   0.420    -6.33e-07    2.64e-07
   emp2digit |   8.60e-09   2.94e-09     2.92   0.004     2.82e-09    1.44e-08
------------------------------------------------------------------------------
        fips |   absorbed                                     (25 categories)
```

DD estimates with covariates, for medium firms, in level

```
Regression with robust standard errors                 Number of obs =    5344
                                                       F( 29,  5149) =37570.01
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7360
                                                       Adj R-squared =  0.7261
                                                       Root MSE      =  .10208
                             (standard errors adjusted for clustering on fips)
```

```
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .0106463   .0075764     1.41   0.160    -.0042065    .0254992
       lagrd |   .0004626   .0003603     1.28   0.199    -.0002438    .0011691
      lagemp |   .8455898    .022534    37.53   0.000     .8014137     .889766
    lagsales |   .0002993   .0001331     2.25   0.025     .0000383    .0005603
  knowcounty |  -1.83e-06   1.70e-06    -1.08   0.281    -5.16e-06    1.50e-06
   emp2digit |   4.65e-08   1.48e-08     3.14   0.002     1.75e-08    7.56e-08
------------------------------------------------------------------------------
     fips |   absorbed                                       (29 categories)
```

DD estimates with covariates, for large firms, in level

```
Regression with robust standard errors                 Number of obs =   18504
                                                       F( 33, 18197) =25812.01
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9850
                                                       Adj R-squared =  0.9847
                                                       Root MSE      =  4.9783
                               (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .4812659   .2700164     1.78   0.075    -.0479916    1.010523
       lagrd |   -.001686   .0031342    -0.54   0.591    -.0078293    .0044574
      lagemp |   .9940136   .0225703    44.04   0.000     .9497736    1.038254
    lagsales |  -.0000117   .0001063    -0.11   0.913    -.0002201    .0001967
  knowcounty |  -.0000126   .0000152    -0.83   0.409    -.0000425    .0000173
   emp2digit |   7.99e-08   3.09e-07     0.26   0.796    -5.26e-07    6.85e-07
------------------------------------------------------------------------------
     fips |   absorbed                                       (28 categories)
```

DD estimates with covariates, for small firms, in log

```
Regression with robust standard errors                 Number of obs =    9642
                                                       F( 26,  9450) = 1219.28
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8099
                                                       Adj R-squared =  0.8061
                                                       Root MSE      =  .43521
                               (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .0276398   .0308067     0.90   0.370     -.032748    .0880275
    loglagrd |   .0166829   .0059204     2.82   0.005     .0050776    .0282882
   loglagemp |   .8724062   .0176038    49.56   0.000     .8378989    .9069136
 loglagsales |   .0263776   .0069131     3.82   0.000     .0128264    .0399288
logknowcou~y |   -.001207    .002937    -0.41   0.681    -.0069642    .0045503
   logemp2di |   .0254538   .0088664     2.87   0.004     .0080737    .0428339
------------------------------------------------------------------------------
     fips |   absorbed                                       (25 categories)
```

DD estimates with covariates, for medium firms, in log

```
Regression with robust standard errors                 Number of obs =    4929
                                                       F( 28,  4747) = 3283.54
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7504
                                                       Adj R-squared =  0.7409
                                                       Root MSE      =  .31513
                               (standard errors adjusted for clustering on fips)
```

logemp	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
tpostdd	.0384611	.0234614	1.64	0.101	-.0075342	.0844564
loglagrd	.0289002	.0099631	2.90	0.004	.0093679	.0484326
loglagemp	.7844979	.0195376	40.15	0.000	.7461951	.8228006
loglagsales	.012699	.0147591	0.86	0.390	-.0162358	.0416337
logknowcou~y	-.0083756	.0043226	-1.94	0.053	-.01685	.0000988
logemp2di	.018672	.0131704	1.42	0.156	-.007148	.044492
fips	absorbed					(29 categories)

DD estimates with covariates, for large firms, in log

```
Regression with robust standard errors                 Number of obs =   17867
                                                       F( 33, 17573) = 6528.60
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9762
                                                       Adj R-squared =  0.9758
                                                       Root MSE      =  .23962
                             (standard errors adjusted for clustering on fips)
```

logemp	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
tpostdd	-.0081106	.0111891	-0.72	0.469	-.0300423	.0138212
loglagrd	.0240153	.0043566	5.51	0.000	.015476	.0325545
loglagemp	.8999936	.0074732	120.43	0.000	.8853453	.9146418
loglagsales	.0431573	.0102914	4.19	0.000	.0229851	.0633295
logknowcou~y	-.0001453	.0009404	-0.15	0.877	-.0019885	.001698
logemp2di	.0008367	.0041563	0.20	0.840	-.0073101	.0089836
fips	absorbed					(28 categories)

A.4.2 The matching estimates for the all industry

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for small firms in the all industry

Variable	Sample	Treated	Controls	Difference
emp_con_d	Unmatched	.009512771	.005617752	.003895019
	ATT	.009512771	.006118873	.003393898

Treatment assignment	support On suppor	Total
Untreated	152	152
Treated	267	267
Total	419	419

```
Bootstrap statistics                             Number of obs    =      1502
                                                 Replications     =        50
```

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf. Interval]		
_bs_1	50	.0033939	-.0035069	.005792	-.0082455	.0150333	(N)

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for small firms in the all industry

Probit estimates — Number of obs = 419; LR chi2(16) = 314.00; Prob > chi2 = 0.0000; Pseudo R2 = 0.5721

Log likelihood = -117.4427

tst	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lagemp	2.150178	2.861908	0.75	0.452	-3.459059	7.759415
lagrd	-.0055014	.0788723	-0.07	0.944	-.1600884	.1490855
lagsales	-.0181998	.0239299	-0.76	0.447	-.0651015	.028702
knowcounty	.0003949	.000102	3.87	0.000	.0001949	.0005948
emp2digit	-1.22e-06	8.32e-07	-1.47	0.142	-2.85e-06	4.10e-07
firmsizeall	-.8382033	2.752092	-0.30	0.761	-6.232204	4.555798
incapita89~n	.0000414	.0000898	0.46	0.645	-.0001347	.0002175
edu1290	-.0138962	.0331645	-0.42	0.675	-.0788975	.0511051
edu1690	-.0973756	.0359112	-2.71	0.007	-.1677602	-.0269909
sbir8690	9.80e-08	1.01e-08	9.68	0.000	7.82e-08	1.18e-07
univ8690	-4.14e-06	5.38e-07	-7.70	0.000	-5.20e-06	-3.09e-06
popchange	-1.964002	.5829092	-3.37	0.001	-3.106483	-.8215214
crime91	.0001765	.0000567	3.11	0.002	.0000653	.0002877
poverty89	-.1510564	.0589245	-2.56	0.010	-.2665462	-.0355666
unemprat90	-.3391901	.1268228	-2.67	0.007	-.5877582	-.090622
mestdensi~87	805.3733	206.4022	3.90	0.000	400.8324	1209.914
_cons	4.401205	3.168372	1.39	0.165	-1.80869	10.6111

Variable	Sample	Treated	Controls	Difference
emp_con_d	Unmatched	.009512771	.005617752	.003895019
	ATT	.014215362	-.014535297	.028750659

Treatment assignment	support: Off suppo	On suppor	Total
Untreated	0	152	152
Treated	200	67	267
Total	200	219	419

Bootstrap statistics — Number of obs = 419; Replications = 50

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf.	Interval]	
_bs_1	50	.0287507	-.0138079	.0243995	-.0202819	.0777832	(N)

Matching estimates with propensity score metrics, with caliper, # of match=3,in level for small firms in the all industry

Variable	Sample	Treated	Controls	Difference
emp_con_d	Unmatched	.009512771	.005617752	.003895019
	ATT	.014215362	-.011883106	.026098468

Treatment assignment	support: Off suppo	On suppor	Total
Untreated	0	152	152
Treated	200	67	267
Total	200	219	419

```
Bootstrap statistics                              Number of obs    =       419
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0260985 -.0128745  .0203575  -.0148114   .0670083   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for medium firms in the all industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .053099438   .057460469  -.004361031
                        ATT | .053099438   .056304912  -.003205474
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |        58 |        58
   Treated |       113 |       113
-----------+-----------+----------
     Total |       171 |       171
Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -.0032055 -.0148345  .0400995  -.0837884   .0773774   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for medium firms in the all industry

```
Probit estimates                                  Number of obs   =        171
                                                  LR chi2(16)     =     115.23
                                                  Prob > chi2     =     0.0000
Log likelihood = -51.906849                       Pseudo R2       =     0.5261
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |  -2.600152   1.567124    -1.66   0.097    -5.671658     .471355
        lagrd |   .0034739   .0349761     0.10   0.921    -.0650781    .0720259
    lagsales |   .0077796   .0089304     0.87   0.384    -.0097235    .0252828
  knowcounty |   .0002967   .0001778     1.67   0.095    -.0000518    .0006452
   emp2digit |   2.69e-07   1.61e-06     0.17   0.867    -2.88e-06    3.42e-06
 firmsizeall |   1.157358   1.933965     0.60   0.550    -2.633143    4.947859
incapita89~n |  -.0000479   .0001355    -0.35   0.723    -.0003135    .0002176
     edu1290 |  -.0524749   .0568333    -0.92   0.356    -.1638662    .0589163
     edu1690 |  -.0123075   .0494614    -0.25   0.803    -.1092501    .0846352
    sbir8690 |   8.65e-08   1.50e-08     5.77   0.000     5.71e-08    1.16e-07
    univ8690 |  -3.83e-06   8.84e-07    -4.34   0.000    -5.56e-06   -2.10e-06
   popchange |  -1.426132    1.25621    -1.14   0.256    -3.888259    1.035995
     crime91 |   .0001278   .0000988     1.29   0.196    -.0000659    .0003215
   poverty89 |  -.2332033   .0934017    -2.50   0.013    -.4162673   -.0501393
  unemprat90 |    .073696   .2229955     0.33   0.741    -.3633671    .5107592
mestdensi~87 |   127.8321   312.6269     0.41   0.683    -484.9054    740.5696
       _cons |   6.750478   5.152991     1.31   0.190    -3.349198    16.85015
------------------------------------------------------------------------------
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .053099438   .057460469  -.004361031
                        ATT |  .15585531   .098903073   .056952237
```

```
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         58 |        58
   Treated |        92         21 |       113
-----------+----------------------+----------
     Total |        92         79 |       171

Bootstrap statistics                              Number of obs    =       171
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0569522 -.0300178  .1382649  -.2209015    .334806   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level for medium firms in the all industry

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .053099438   .057460469  -.004361031
                        ATT |  .15585531   .100472521   .055382789
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         58 |        58
   Treated |        92         21 |       113
-----------+----------------------+----------
     Total |        92         79 |       171

Bootstrap statistics                              Number of obs    =       171
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0553828   .013923  .1637243  -.2736335   .3843991   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for large firms in the all industry

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | 1.17487995  -.866821045   2.04170099
                        ATT | 1.17487995  -.223488949    1.3983689
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       209 |       209
   Treated |       398 |       398
-----------+-----------+----------
     Total |       607 |       607

Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  1.398369 -.3478274  .9175118    -.44544   3.242178   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for large firms in the all industry

Probit estimates

	Number of obs	=	607
	LR chi2(16)	=	348.19
	Prob > chi2	=	0.0000
Log likelihood = -216.72667	Pseudo R2	=	0.4455

tst	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lagemp	.0142047	.014296	0.99	0.320	-.0138149	.0422244
lagrd	.0010097	.0008635	1.17	0.242	-.0006827	.0027021
lagsales	-.0001244	.0000778	-1.60	0.110	-.000277	.0000281
knowcounty	.0003664	.000072	5.09	0.000	.0002252	.0005076
emp2digit	2.53e-07	8.89e-07	0.29	0.776	-1.49e-06	1.99e-06
firmsizeall	-.0095851	.0121003	-0.79	0.428	-.0333013	.014131
incapita89~n	-.0000807	.0000632	-1.28	0.202	-.0002046	.0000432
edu1290	.0221015	.0260499	0.85	0.396	-.0289554	.0731584
edu1690	-.0558794	.0268634	-2.08	0.038	-.1085306	-.0032282
sbir8690	7.66e-08	7.57e-09	10.12	0.000	6.17e-08	9.14e-08
univ8690	-2.99e-06	3.95e-07	-7.56	0.000	-3.76e-06	-2.21e-06
popchange	-2.495284	.5659513	-4.41	0.000	-3.604528	-1.38604
crime91	.0001134	.0000351	3.23	0.001	.0000446	.0001823
poverty89	-.131651	.0419578	-3.14	0.002	-.2138868	-.0494153
unemprat90	-.3080915	.0971718	-3.17	0.002	-.4985447	-.1176384
mestdensi~87	415.159	129.8188	3.20	0.001	160.7188	669.5991
_cons	2.85288	2.208855	1.29	0.197	-1.476396	7.182156

Variable	Sample	Treated	Controls	Difference
emp_con_d	Unmatched	1.17487995	-.866821045	2.04170099
	ATT	.990821233	3.10204734	-2.1112261

Treatment assignment	support Off suppo	On suppor	Total
Untreated	0	209	209
Treated	222	176	398
Total	222	385	607

Bootstrap statistics

	Number of obs	=	607
	Replications	=	50

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf.	Interval]	
_bs_1	50	-2.111226	3.268957	2.630678	-7.397771	3.175319	(N)

Matching estimates with propensity score metrics, with caliper, # of match=3,in level for large firms in the all industry

Variable	Sample	Treated	Controls	Difference
emp_con_d	Unmatched	1.17487995	-.866821045	2.04170099
	ATT	.990821233	2.27696663	-1.28614539

Treatment assignment	support Off suppo	On suppor	Total
Untreated	0	209	209
Treated	222	176	398
Total	222	385	607

```
Bootstrap statistics                              Number of obs    =       607
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -1.286145  2.552356  2.508724  -6.327614   3.755323   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log for small firms in the all industry

```
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
        logemp_d  Unmatched | .091741677    .01750616   .074235517
                        ATT | .091741677   .081627201   .010114477
----------------------------+-------------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       152 |       152
   Treated |       267 |       267
-----------+-----------+----------
     Total |       419 |       419
Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0101145 -.0564545  .1240569  -.2391872   .2594161   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in log for small firms in the all industry

```
Probit estimates                                  Number of obs   =        419
                                                  LR chi2(16)     =     314.00
                                                  Prob > chi2     =     0.0000
Log likelihood =  -117.4427                       Pseudo R2       =     0.5721
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |   2.150178   2.861908     0.75   0.452    -3.459059    7.759415
       lagrd |  -.0055014   .0788723    -0.07   0.944    -.1600884    .1490855
    lagsales |  -.0181998   .0239299    -0.76   0.447    -.0651015     .028702
  knowcounty |   .0003949    .000102     3.87   0.000     .0001949    .0005948
   emp2digit |  -1.22e-06   8.32e-07    -1.47   0.142    -2.85e-06    4.10e-07
 firmsizeall |  -.8382033   2.752092    -0.30   0.761    -6.232204    4.555798
incapita89~n |   .0000414   .0000898     0.46   0.645    -.0001347    .0002175
     edu1290 |  -.0138962   .0331645    -0.42   0.675    -.0788975    .0511051
     edu1690 |  -.0973756   .0359112    -2.71   0.007    -.1677602   -.0269909
    sbir8690 |   9.80e-08   1.01e-08     9.68   0.000     7.82e-08    1.18e-07
    univ8690 |  -4.14e-06   5.38e-07    -7.70   0.000    -5.20e-06   -3.09e-06
   popchange |  -1.964002   .5829092    -3.37   0.001    -3.106483   -.8215214
     crime91 |   .0001765   .0000567     3.11   0.002     .0000653    .0002877
   poverty89 |  -.1510564   .0589245    -2.56   0.010    -.2665462   -.0355666
  unemprat90 |  -.3391901   .1268228    -2.67   0.007    -.5877582    -.090622
mestdensi~87 |   805.3733   206.4022     3.90   0.000     400.8324    1209.914
       _cons |   4.401205   3.168372     1.39   0.165     -1.80869     10.6111
------------------------------------------------------------------------------
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
        logemp_d  Unmatched | .091741677    .01750616   .074235517
                        ATT | .180439877  -.273160058   .453599935
```

```
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        152 |       152
   Treated |       218         49 |       267
-----------+----------------------+----------
     Total |       218        201 |       419
Bootstrap statistics                              Number of obs    =       419
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .4535999 -.3191171  .3988202  -.3478592   1.255059   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in log for small firms in the all industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .091741677    .01750616   .074235517
                        ATT | .180439877  -.245582308   .426022185
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        152 |       152
   Treated |       218         49 |       267
-----------+----------------------+----------
     Total |       218        201 |       419

Bootstrap statistics                              Number of obs    =       419
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .4260222 -.2504019  .3266939  -.2304938   1.082538   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log for medium firms in the all industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .209645332   .247759832  -.038114501
                        ATT | .209645332     .1823639   .027281432
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |        58 |        58
   Treated |       113 |       113
-----------+-----------+----------
    Total |       171 |       171

Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0272814 -.0007406   .157162  -.2885475   .3431104   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in log for medium firms in the all industry

Probit estimates	Number of obs	=	171
	LR chi2(16)	=	115.23
	Prob > chi2	=	0.0000
Log likelihood = -51.906849	Pseudo R2	=	0.5261

tst	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lagemp	-2.600152	1.567124	-1.66	0.097	-5.671658	.471355
lagrd	.0034739	.0349761	0.10	0.921	-.0650781	.0720259
lagsales	.0077796	.0089304	0.87	0.384	-.0097235	.0252828
knowcounty	.0002967	.0001778	1.67	0.095	-.0000518	.0006452
emp2digit	2.69e-07	1.61e-06	0.17	0.867	-2.88e-06	3.42e-06
firmsizeall	1.157358	1.933965	0.60	0.550	-2.633143	4.947859
incapita89~n	-.0000479	.0001355	-0.35	0.723	-.0003135	.0002176
edu1290	-.0524749	.0568333	-0.92	0.356	-.1638662	.0589163
edu1690	-.0123075	.0494614	-0.25	0.803	-.1092501	.0846352
sbir8690	8.65e-08	1.50e-08	5.77	0.000	5.71e-08	1.16e-07
univ8690	-3.83e-06	8.84e-07	-4.34	0.000	-5.56e-06	-2.10e-06
popchange	-1.426132	1.25621	-1.14	0.256	-3.888259	1.035995
crime91	.0001278	.0000988	1.29	0.196	-.0000659	.0003215
poverty89	-.2332033	.0934017	-2.50	0.013	-.4162673	-.0501393
unemprat90	.073696	.2229955	0.33	0.741	-.3633671	.5107592
mestdensi~87	127.8321	312.6269	0.41	0.683	-484.9054	740.5696
_cons	6.750478	5.152991	1.31	0.190	-3.349198	16.85015

Variable	Sample	Treated	Controls	Difference
logemp_d	Unmatched	.209645332	.247759832	-.038114501
	ATT	.479214767	.482637098	-.003422331

Treatment assignment	support Off suppo	On suppor	Total
Untreated	0	58	58
Treated	92	21	113
Total	92	79	171

Bootstrap statistics	Number of obs	=	171
	Replications	=	50

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf.	Interval]	
_bs_1	50	-.0034223	.0966343	.3667888	-.740512	.7336674	(N)

Matching estimates with propensity score metrics, with caliper, # of match=3,in log for medium firms in the all industry

Variable	Sample	Treated	Controls	Difference
logemp_d	Unmatched	.209645332	.247759832	-.038114501
	ATT	.479214767	.496890767	-.017676

Treatment assignment	support Off suppo	On suppor	Total
Untreated	0	58	58
Treated	92	21	113
Total	92	79	171

```
Bootstrap statistics                              Number of obs    =       171
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  -.017676  .1896629  .4479148  -.9177945   .8824425   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log for large firms in the all industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .452874222   .143767787   .309106435
                        ATT | .452874222   .188966861   .263907361
----------------------------+-----------------------------------
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       209 |       209
   Treated |       398 |       398
-----------+-----------+----------
     Total |       607 |       607
Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .2639074 -.0230178  .1053796   .0521392   .4756756   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in log for large firms in the all industry

```
Probit estimates                                  Number of obs   =        607
                                                  LR chi2(16)     =     348.19
                                                  Prob > chi2     =     0.0000
Log likelihood = -216.72667                       Pseudo R2       =     0.4455
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |   .0142047    .014296     0.99   0.320    -.0138149    .0422244
       lagrd |   .0010097   .0008635     1.17   0.242    -.0006827    .0027021
    lagsales |  -.0001244   .0000778    -1.60   0.110     -.000277    .0000281
  knowcounty |   .0003664    .000072     5.09   0.000     .0002252    .0005076
   emp2digit |   2.53e-07   8.89e-07     0.29   0.776    -1.49e-06    1.99e-06
 firmsizeall |  -.0095851   .0121003    -0.79   0.428    -.0333013     .014131
incapita89~n |  -.0000807   .0000632    -1.28   0.202    -.0002046    .0000432
     edu1290 |   .0221015   .0260499     0.85   0.396    -.0289554    .0731584
     edu1690 |  -.0558794   .0268634    -2.08   0.038    -.1085306   -.0032282
    sbir8690 |   7.66e-08   7.57e-09    10.12   0.000     6.17e-08    9.14e-08
    univ8690 |  -2.99e-06   3.95e-07    -7.56   0.000    -3.76e-06   -2.21e-06
   popchange |  -2.495284   .5659513    -4.41   0.000    -3.604528    -1.38604
     crime91 |   .0001134   .0000351     3.23   0.001     .0000446    .0001823
   poverty89 |   -.131651   .0419578    -3.14   0.002    -.2138868   -.0494153
  unemprat90 |  -.3080915   .0971718    -3.17   0.002    -.4985447   -.1176384
mestdensi~87 |    415.159   129.8188     3.20   0.001     160.7188    669.5991
       _cons |    2.85288   2.208855     1.29   0.197    -1.476396    7.182156
------------------------------------------------------------------------------
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .452874222   .143767787   .309106435
                        ATT | .204410415   .269378352  -.064967937
```

```
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        209 |       209
   Treated |       267        131 |       398
-----------+----------------------+----------
     Total |       267        340 |       607

Bootstrap statistics                              Number of obs    =       607
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -.0649679  .0713315  .2057507  -.4784395   .3485036   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in log for large firms in the all industry

```
----------------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .452874222   .143767787   .309106435
                        ATT | .204410415   .299369935   -.09495952
----------------------------+-----------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        209 |       209
   Treated |       267        131 |       398
-----------+----------------------+----------
     Total |       267        340 |       607

Bootstrap statistics                              Number of obs    =       607
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -.0949595  .1645817  .1573439  -.4111539   .2212349   (N)
------------------------------------------------------------------------------
```

A.4.3 The DD/DDD estimates for the high-technology industry

DD estimates with covariates, for small firms, in level

```
Regression with robust standard errors                 Number of obs =    8951
                                                       F( 25,  8841) = 4923.06
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7530
                                                       Adj R-squared =  0.7500
                                                       Root MSE      =  .03805
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostdd |   .0008062   .0031653     0.25   0.799    -.0053986    .0070109
        lagrd |   .0005098   .0003042     1.68   0.094    -.0000864     .001106
       lagemp |   .8384675    .031422    26.68   0.000     .7768732    .9000619
     lagsales |   .0004763   .0000861     5.53   0.000     .0003076     .000645
   knowcounty |  -1.46e-07   1.35e-07    -1.09   0.278    -4.10e-07    1.18e-07
```

```
   emp2digit |   6.21e-09   3.47e-09     1.79   0.074    -5.93e-10    1.30e-08
------------------------------------------------------------------------------
        fips |   absorbed                                      (25 categories)
```

DD estimates with covariates, for medium firms, in level

```
Regression with robust standard errors                 Number of obs =    3523
                                                       F( 28,  3414) = 4864.75
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7363
                                                       Adj R-squared =  0.7280
                                                       Root MSE      =  .09976
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   .0249219   .0083121     3.00   0.003     .0086247     .041219
       lagrd |   .0003998    .000389     1.03   0.304    -.0003628    .0011625
      lagemp |   .8209554   .0231309    35.49   0.000     .7756036    .8663073
    lagsales |   .0004661   .0001327     3.51   0.000      .000206    .0007263
  knowcounty |  -3.11e-06   1.58e-06    -1.97   0.049    -6.20e-06   -7.70e-09
   emp2digit |   5.52e-08   1.64e-08     3.36   0.001     2.30e-08    8.75e-08
------------------------------------------------------------------------------
        fips |   absorbed                                      (29 categories)
```

DD estimates with covariates, for large firms, in level

```
Regression with robust standard errors                 Number of obs =    8518
                                                       F( 27,  8395) =  745.76
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9911
                                                       Adj R-squared =  0.9910
                                                       Root MSE      =  4.8916
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   .4752388   .2445443     1.94   0.052    -.0041283     .954606
       lagrd |  -.0041694   .0031754    -1.31   0.189     -.010394    .0020553
      lagemp |   1.024491    .017549    58.38   0.000     .9900902    1.058891
    lagsales |  -.0001093   .0000956    -1.14   0.253    -.0002966     .000078
  knowcounty |   .0000286   .0000316     0.91   0.365    -.0000334    .0000906
   emp2digit |  -6.23e-08   3.74e-07    -0.17   0.868    -7.96e-07    6.71e-07
------------------------------------------------------------------------------
        fips |   absorbed                                      (28 categories)
```

DDD estimates with covariates, for small firms, in level

```
Regression with robust standard errors                 Number of obs =   11566
                                                       F( 26, 11350) =14197.32
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7539
                                                       Adj R-squared =  0.7492
                                                       Root MSE      =  .03844
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostddd |  -.0015483   .0029987    -0.52   0.606    -.0074262    .0043296
       lagrd |   .0005194   .0003206     1.62   0.105    -.0001091    .0011479
      lagemp |   .8346828   .0298075    28.00   0.000     .7762549    .8931108
    lagsales |   .0005002   .0000705     7.10   0.000      .000362    .0006384
```

```
  knowcounty |  -1.81e-07   2.36e-07    -0.77   0.443    -6.44e-07    2.82e-07
   emp2digit |   8.65e-09   2.96e-09     2.93   0.003     2.86e-09    1.44e-08
------------------------------------------------------------------------------
        fips |   absorbed                                     (25 categories)
```

DDD estimates with covariates, for medium firms, in level

```
Regression with robust standard errors                 Number of obs =    5344
                                                       F( 29,  5146) =11425.57
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7367
                                                       Adj R-squared =  0.7266
                                                       Root MSE      =  .10198
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostddd |    .029666   .0145323     2.04   0.041     .0011766    .0581554
       lagrd |   .0004842   .0003666     1.32   0.187    -.0002345    .0012028
      lagemp |   .8424941   .0212153    39.71   0.000      .800903    .8840851
    lagsales |   .0003136   .0001272     2.47   0.014     .0000643     .000563
  knowcounty |  -1.91e-06   1.67e-06    -1.15   0.251    -5.18e-06    1.36e-06
   emp2digit |   4.80e-08   1.55e-08     3.09   0.002     1.76e-08    7.85e-08
------------------------------------------------------------------------------
        fips |   absorbed                                     (29 categories)
```

DDD estimates with covariates, for large firms, in level

```
Regression with robust standard errors                 Number of obs =   18504
                                                       F( 33, 18194) =29819.28
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9850
                                                       Adj R-squared =  0.9847
                                                       Root MSE      =  4.9784
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
         emp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostddd |    .308865   .3276771     0.94   0.346     -.333413    .9511431
       lagrd |   -.001663   .0031307    -0.53   0.595    -.0077995    .0044735
      lagemp |    .993625   .0224629    44.23   0.000     .9495956    1.037654
    lagsales |  -9.07e-06   .0001053    -0.09   0.931    -.0002154    .0001973
  knowcounty |  -.0000122    .000015    -0.81   0.416    -.0000417    .0000172
   emp2digit |   9.44e-08   3.21e-07     0.29   0.769    -5.35e-07    7.24e-07
------------------------------------------------------------------------------
        fips |   absorbed                                     (28 categories)
```

DD estimates with covariates, for small firms, in log

```
Regression with robust standard errors                 Number of obs =    7616
                                                       F( 25,  7507) =  384.99
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8063
                                                       Adj R-squared =  0.8035
                                                       Root MSE      =  .43731
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |    .021578   .0291067     0.74   0.459    -.0354793    .0786354
    loglagrd |   .0144564   .0078201     1.85   0.065    -.0008731    .0297859
   loglagemp |   .8812699   .0204526    43.09   0.000     .8411771    .9213627
 loglagsales |   .0240007   .0076111     3.15   0.002     .0090808    .0389206
logknowcou~y |  -.0014757   .0019056    -0.77   0.439    -.0052113    .0022599
```

```
   logemp2di |   .0268216   .0104476     2.57   0.010     .0063414    .0473018
------------------------------------------------------------------------------
        fips |   absorbed                                     (25 categories)
```

DD estimates with covariates, for medium firms, in log

```
Regression with robust standard errors                 Number of obs =    3381
                                                       F( 28,  3273) =  144.85
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7453
                                                       Adj R-squared =  0.7370
                                                       Root MSE      =  .31068
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   .0422617   .0280717     1.51   0.132    -.0127781    .0973016
    loglagrd |    .021329   .0101052     2.11   0.035     .0015158    .0411422
   loglagemp |   .7804456   .0178116    43.82   0.000     .7455226    .8153686
 loglagsales |   .0086322   .0114046     0.76   0.449    -.0137287     .030993
logknowcou~y |  -.0085807    .004198    -2.04   0.041    -.0168116   -.0003498
   logemp2di |   .0202078   .0122104     1.65   0.098     -.003733    .0441486
------------------------------------------------------------------------------
        fips |   absorbed                                     (29 categories)
```

DD estimates with covariates, for large firms, in log

```
Regression with robust standard errors                 Number of obs =    8447
                                                       F( 27,  8324) =  221.00
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9751
                                                       Adj R-squared =  0.9747
                                                       Root MSE      =  .25248
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tpostdd |   -.004969   .0170346    -0.29   0.771     -.038361     .028423
    loglagrd |   .0297665     .00632     4.71   0.000     .0173777    .0421553
   loglagemp |   .9002934   .0102105    88.17   0.000     .8802783    .9203085
 loglagsales |    .030627   .0143286     2.14   0.033     .0025394    .0587147
logknowcou~y |   .0013482     .00175     0.77   0.441    -.0020822    .0047785
   logemp2di |   .0019081   .0089786     0.21   0.832    -.0156923    .0195084
------------------------------------------------------------------------------
        fips |   absorbed                                     (28 categories)
```

DDD estimates with covariates, for small firms, in log

```
Regression with robust standard errors                 Number of obs =    9642
                                                       F( 26,  9447) =  774.05
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8100
                                                       Adj R-squared =  0.8061
                                                       Root MSE      =   .4352
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostddd |  -.0238027   .0468425    -0.51   0.611     -.115624    .0680187
    loglagrd |   .0163391   .0057587     2.84   0.005     .0050508    .0276275
   loglagemp |   .8723549   .0178074    48.99   0.000     .8374485    .9072613
```

```
 loglagsales |   .0263236   .0068755     3.83   0.000     .0128461     .039801
logknowcou~y |  -.0012367   .0028999    -0.43   0.670    -.0069212    .0044477
   logemp2di |    .025745    .008949     2.88   0.004      .008203    .0432871
------------------------------------------------------------------------------
        fips |   absorbed                                      (25 categories)
```

DDD estimates with covariates, for medium firms, in log

```
Regression with robust standard errors                 Number of obs =    4929
                                                       F( 28,  4744) =19267.25
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7504
                                                       Adj R-squared =  0.7407
                                                       Root MSE      =  .31522
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostddd |  -.0056377   .0445549    -0.13   0.899     -.092986    .0817107
    loglagrd |   .0288985   .0098953     2.92   0.004      .009499    .0482979
   loglagemp |   .7845377   .0199202    39.38   0.000     .7454849    .8235906
 loglagsales |   .0126629    .015062     0.84   0.401    -.0168655    .0421914
logknowcou~y |  -.0083048   .0042956    -1.93   0.053    -.0167261    .0001166
   logemp2di |   .0187063   .0127138     1.47   0.141    -.0062187    .0436313
------------------------------------------------------------------------------
        fips |   absorbed                                      (29 categories)
```

DDD estimates with covariates, for large firms, in log

```
Regression with robust standard errors                 Number of obs =   17867
                                                       F( 33, 17570) = 6.5e+05
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9762
                                                       Adj R-squared =  0.9758
                                                       Root MSE      =  .23952
                             (standard errors adjusted for clustering on fips)
------------------------------------------------------------------------------
             |               Robust
      logemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    tpostddd |   .0123574   .0211465     0.58   0.559    -.0290918    .0538066
    loglagrd |   .0239883    .004432     5.41   0.000     .0153012    .0326754
   loglagemp |   .8979633   .0074861   119.95   0.000     .8832898    .9126369
 loglagsales |   .0454631   .0101627     4.47   0.000     .0255432     .065383
logknowcou~y |  -.0001781    .000936    -0.19   0.849    -.0020128    .0016566
   logemp2di |   .0012285   .0044343     0.28   0.782    -.0074632    .0099202
------------------------------------------------------------------------------
        fips |   absorbed                                      (28 categories)
```

A.4.4 The matching estimates for the high-technology industry

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for small firms in the high-technology industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .010284041   .008897832   .001386209
                        ATT | .010284041   .007755364   .002528677
```

```
 Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       116 |       116
   Treated |       207 |       207
-----------+-----------+----------
     Total |       323 |       323

Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0025287 -.0017465  .0065536  -.0106413   .0156986   (N)
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for small firms in the high-technology industry

```
Probit estimates                                  Number of obs   =        323
                                                  LR chi2(16)     =     241.36
                                                  Prob > chi2     =     0.0000
Log likelihood = -90.214572                       Pseudo R2       =     0.5722
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |   3.863153   3.163822     1.22   0.222    -2.337824    10.06413
       lagrd |  -.0284166   .0818487    -0.35   0.728     -.188837    .1320039
    lagsales |  -.0118988   .0261235    -0.46   0.649    -.0630998    .0393022
  knowcounty |   .0002697    .000119     2.27   0.023     .0000364    .0005029
   emp2digit |  -1.59e-06   9.94e-07    -1.60   0.110    -3.54e-06    3.60e-07
 firmsizeall |  -2.941897   3.096531    -0.95   0.342    -9.010986    3.127191
incapita89~n |   .0000385   .0001013     0.38   0.704      -.00016    .0002371
     edu1290 |  -.0687917   .0439746    -1.56   0.118    -.1549802    .0173969
     edu1690 |  -.0666159   .0404773    -1.65   0.100    -.1459501    .0127182
    sbir8690 |   9.80e-08   1.19e-08     8.26   0.000     7.48e-08    1.21e-07
    univ8690 |  -3.97e-06   6.15e-07    -6.46   0.000    -5.18e-06   -2.77e-06
   popchange |   -1.24056   .6394194    -1.94   0.052    -2.493799    .0126785
     crime91 |   .0002674   .0000719     3.72   0.000     .0001264    .0004083
   poverty89 |  -.1978938   .0692961    -2.86   0.004    -.3337117   -.0620759
  unemprat90 |  -.3688029   .1601804    -2.30   0.021    -.6827508   -.0548551
mestdensi~87 |   1144.964   273.3384     4.19   0.000     609.2309    1680.698
       _cons |   7.391084   4.078062     1.81   0.070     -.601771    15.38394
------------------------------------------------------------------------------
----------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------------
       emp_con_d  Unmatched | .010284041   .008897832   .001386209
                        ATT | .007108888   .003595714   .003513174
----------------------------+-----------------------------------------
 Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        116 |       116
   Treated |       156         51 |       207
-----------+----------------------+----------
     Total |       156        167 |       323

Bootstrap statistics                              Number of obs    =       323
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0035132   .007234  .0230068  -.0427207   .0497471   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level for small firms in the high-technology industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .010284041   .008897832   .001386209
                        ATT | .007108888    -.001299   .008407888
----------------------------+-----------------------------------
Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        116 |       116
   Treated |       156         51 |       207
-----------+----------------------+----------
     Total |       156        167 |       323
Bootstrap statistics                              Number of obs    =       323
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0084079  .0009556  .0223881  -.0365827   .0533985   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for medium firms in the high-technology industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .033355432    .06577526  -.032419828
                        ATT | .033355432   .090043764  -.056688332
----------------------------+-----------------------------------
Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |        31 |        31
   Treated |        84 |        84
-----------+-----------+----------
     Total |       115 |       115
Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -.0566883  .0097405  .0643912  -.1860872   .0727105   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for medium firms in the high-technology industry

```
Probit estimates                                  Number of obs   =        115
                                                  LR chi2(16)     =      82.86
                                                  Prob > chi2     =     0.0000
Log likelihood = -25.592608                       Pseudo R2       =     0.6182
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |  -6.098704    2.73409    -2.23   0.026    -11.45742   -.7399854
       lagrd |   .0172801   .0438662     0.39   0.694    -.0686961    .1032562
    lagsales |   .0177364   .0148182     1.20   0.231    -.0113067    .0467796
  knowcounty |   .0007203   .0003265     2.21   0.027     .0000804    .0013602
   emp2digit |   9.75e-07   2.72e-06     0.36   0.720    -4.36e-06    6.31e-06
 firmsizeall |   .3543142   3.320105     0.11   0.915    -6.152973    6.861601
incapita89~n |  -.0000457   .0001897    -0.24   0.810    -.0004176    .0003262
     edu1290 |   .0289167   .0849877     0.34   0.734    -.1376561    .1954894
```

```
     edu1690 |  -.1338045   .0837522    -1.60   0.110    -.2979558    .0303469
    sbir8690 |   1.30e-07   3.21e-08     4.06   0.000     6.75e-08    1.93e-07
    univ8690 |  -6.04e-06   1.82e-06    -3.31   0.001    -9.62e-06   -2.47e-06
   popchange |  -1.709016   1.767218    -0.97   0.334    -5.172699    1.754667
     crime91 |   .0001132   .0001592     0.71   0.477    -.0001989    .0004252
   poverty89 |  -.2092254   .1329805    -1.57   0.116    -.4698624    .0514116
  unemprat90 |  -.4597339   .4257164    -1.08   0.280    -1.294123    .3746548
mestdensi~87 |   238.3653   507.4357     0.47   0.639    -756.1904    1232.921
       _cons |   6.926791    8.31631     0.83   0.405    -9.372878    23.22646
------------------------------------------------------------------------------
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .033355432    .06577526  -.032419828
                        ATT | .140920303   .000223116   .140697187
----------------------------+-----------------------------------
Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         31 |        31
   Treated |        79          5 |        84
-----------+----------------------+----------
     Total |        79         36 |       115

Bootstrap statistics                              Number of obs    =       115
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    29  .1406972 -.1661403  .1630262  -.1932469   .4746413  (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level for medium firms in the high-technology industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | .033355432    .06577526  -.032419828
                        ATT | .140920303      .036371   .104549303
----------------------------+-----------------------------------
Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         31 |        31
   Treated |        79          5 |        84
-----------+----------------------+----------
     Total |        79         36 |       115

Bootstrap statistics                              Number of obs    =       115
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    27  .1045493  -.104078  .1822786  -.2701297   .4792283  (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in level for large firms in the high-technology industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
       emp_con_d  Unmatched | 1.60123333  -2.19961966     3.800853
                        ATT | 1.60123333   -.44786748   2.04910081
----------------------------+-----------------------------------
```

```
Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |        86 |        86
   Treated |       225 |       225
-----------+-----------+----------
     Total |       311 |       311

Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  2.049101 -.2492722  1.023192  -.0070813   4.105283   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in level for large firms in the high-technology industry

```
Probit estimates                                  Number of obs   =        311
                                                  LR chi2(16)     =     209.12
                                                  Prob > chi2     =     0.0000
Log likelihood = -78.818508                       Pseudo R2       =     0.5702
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |   .0443093   .0311334     1.42   0.155     -.016711    .1053297
       lagrd |   .0000444   .0014689     0.03   0.976    -.0028346    .0029235
    lagsales |  -.0001243   .0001136    -1.09   0.274     -.000347    .0000984
  knowcounty |    .000328   .0001194     2.75   0.006      .000094     .000562
   emp2digit |   2.58e-07   1.39e-06     0.19   0.853    -2.47e-06    2.99e-06
 firmsizeall |  -.0365133   .0305031    -1.20   0.231    -.0962982    .0232716
incapita89~n |  -.0000719   .0001016    -0.71   0.479     -.000271    .0001271
     edu1290 |   .0282586   .0429387     0.66   0.510    -.0558996    .1124168
     edu1690 |  -.0818808   .0413252    -1.98   0.048    -.1628766    -.000885
    sbir8690 |   8.70e-08   1.23e-08     7.07   0.000     6.29e-08    1.11e-07
    univ8690 |  -3.62e-06   6.52e-07    -5.55   0.000    -4.90e-06   -2.34e-06
   popchange |  -2.266128   .7673651    -2.95   0.003    -3.770136   -.7621199
     crime91 |  -.0000603   .0000616    -0.98   0.328    -.0001811    .0000605
   poverty89 |  -.0852375   .0691714    -1.23   0.218    -.2208109    .0503359
  unemprat90 |   -.256056   .1674963    -1.53   0.126    -.5843426    .0722307
mestdensi~87 |   321.7771   217.7516     1.48   0.139    -105.0082    748.5624
       _cons |   3.664107   3.980253     0.92   0.357    -4.137045    11.46526
------------------------------------------------------------------------------
------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-------------------------------------
       emp_con_d  Unmatched | 1.60123333  -2.19961966     3.800853
                        ATT | 2.21975718   .288248063   1.93150912
----------------------------+-------------------------------------
Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         86 |        86
   Treated |       160         65 |       225
-----------+----------------------+----------
     Total |       160        151 |       311

Bootstrap statistics                              Number of obs    =       311
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  1.931509  .6733395  2.887264  -3.870665   7.733684   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in level
for large firms in the high-technology industry

```
-----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+------------------------------------
       emp_con_d  Unmatched | 1.60123333  -2.19961966     3.800853
                        ATT | 2.21975718   .368269714   1.85148747
----------------------------+------------------------------------
Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         86 |        86
   Treated |       160         65 |       225
-----------+----------------------+----------
     Total |       160        151 |       311
Bootstrap statistics                               Number of obs    =       311
                                                   Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  1.851488 -.1948492  2.671472  -3.517037   7.220012   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log
for small firms in the high-technology industry

```
-----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+------------------------------------
        logemp_d  Unmatched | .109016408   .065134551   .043881858
                        ATT | .109016408    .07502672   .033989688
----------------------------+------------------------------------
Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |       116 |       116
   Treated |       207 |       207
-----------+-----------+----------
     Total |       323 |       323
Bootstrap statistics                               Number of obs    =      1502
                                                   Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0339897 -.0305034   .129436  -.2261218   .2941012   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in log
for small firms in the high-technology industry

```
Probit estimates                                   Number of obs   =        323
                                                   LR chi2(16)     =     241.36
                                                   Prob > chi2     =     0.0000
Log likelihood = -90.214572                        Pseudo R2       =     0.5722
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |   3.863153   3.163822     1.22   0.222    -2.337824    10.06413
       lagrd |  -.0284166   .0818487    -0.35   0.728     -.188837    .1320039
    lagsales |  -.0118988   .0261235    -0.46   0.649    -.0630998    .0393022
  knowcounty |   .0002697    .000119     2.27   0.023     .0000364    .0005029
   emp2digit |  -1.59e-06   9.94e-07    -1.60   0.110    -3.54e-06    3.60e-07
 firmsizeall |  -2.941897   3.096531    -0.95   0.342    -9.010986    3.127191
incapita89~n |   .0000385   .0001013     0.38   0.704      -.00016    .0002371
     edu1290 |  -.0687917   .0439746    -1.56   0.118    -.1549802    .0173969
```

```
     edu1690 |  -.0666159   .0404773    -1.65   0.100    -.1459501    .0127182
    sbir8690 |   9.80e-08   1.19e-08     8.26   0.000     7.48e-08    1.21e-07
    univ8690 |  -3.97e-06   6.15e-07    -6.46   0.000    -5.18e-06   -2.77e-06
   popchange |   -1.24056   .6394194    -1.94   0.052    -2.493799    .0126785
     crime91 |   .0002674   .0000719     3.72   0.000     .0001264    .0004083
   poverty89 |  -.1978938   .0692961    -2.86   0.004    -.3337117   -.0620759
  unemprat90 |  -.3688029   .1601804    -2.30   0.021    -.6827508   -.0548551
mestdensi~87 |   1144.964   273.3384     4.19   0.000     609.2309    1680.698
       _cons |   7.391084   4.078062     1.81   0.070     -.601771    15.38394
------------------------------------------------------------------------------
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .109016408   .065134551   .043881858
                        ATT | .231563556  -.025799948   .257363504
Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        116 |       116
   Treated |       173         34 |       207
-----------+----------------------+----------
     Total |       173        150 |       323

Bootstrap statistics                              Number of obs    =       323
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .2573635 -.0658984  .3606778  -.4674456   .9821726   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in log for small firms in the high-technology industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .109016408   .065134551   .043881858
                        ATT | .231563556  -.021047452   .252611007
----------------------------+-----------------------------------
Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0        116 |       116
   Treated |       173         34 |       207
-----------+----------------------+----------
     Total |       173        150 |       323

Bootstrap statistics                              Number of obs    =       323
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50   .252611 -.0810262  .3078177  -.3659718   .8711938   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log for medium firms in the high-technology industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .150486961   .308218881   -.15773192
                        ATT | .150486961   .341417156  -.190930195
----------------------------+-----------------------------------
```

```
Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |        31 |        31
   Treated |        84 |        84
-----------+-----------+----------
     Total |       115 |       115

Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50 -.1909302  .0806039  .2057911  -.6044829   .2226225   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in log for medium firms in the high-technology industry

```
Probit estimates                                  Number of obs   =        115
                                                  LR chi2(16)     =      82.86
                                                  Prob > chi2     =     0.0000
Log likelihood = -25.592608                       Pseudo R2       =     0.6182
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |  -6.098704    2.73409    -2.23   0.026    -11.45742   -.7399854
       lagrd |   .0172801   .0438662     0.39   0.694    -.0686961    .1032562
    lagsales |   .0177364   .0148182     1.20   0.231    -.0113067    .0467796
  knowcounty |   .0007203   .0003265     2.21   0.027     .0000804    .0013602
   emp2digit |   9.75e-07   2.72e-06     0.36   0.720    -4.36e-06    6.31e-06
 firmsizeall |   .3543142   3.320105     0.11   0.915    -6.152973    6.861601
incapita89~n |  -.0000457   .0001897    -0.24   0.810    -.0004176    .0003262
     edu1290 |   .0289167   .0849877     0.34   0.734    -.1376561    .1954894
     edu1690 |  -.1338045   .0837522    -1.60   0.110    -.2979558    .0303469
    sbir8690 |   1.30e-07   3.21e-08     4.06   0.000     6.75e-08    1.93e-07
    univ8690 |  -6.04e-06   1.82e-06    -3.31   0.001    -9.62e-06   -2.47e-06
   popchange |  -1.709016   1.767218    -0.97   0.334    -5.172699    1.754667
     crime91 |   .0001132   .0001592     0.71   0.477    -.0001989    .0004252
   poverty89 |  -.2092254   .1329805    -1.57   0.116    -.4698624    .0514116
  unemprat90 |  -.4597339   .4257164    -1.08   0.280    -1.294123    .3746548
mestdensi~87 |   238.3653   507.4357     0.47   0.639    -756.1904    1232.921
       _cons |   6.926791    8.31631     0.83   0.405    -9.372878    23.22646
------------------------------------------------------------------------------
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .150486961   .308218881   -.15773192
                        ATT |  .95114323   .566504126   .384639104
----------------------------+-----------------------------------
Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         31 |        31
   Treated |        81          3 |        84
-----------+----------------------+----------
     Total |        81         34 |       115

Bootstrap statistics                              Number of obs    =       115
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    24  .3846391   -.52175  .6037659  -.8643459   1.633624   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in log for medium firms in the high-technology industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .150486961   .308218881   -.15773192
                        ATT |  .95114323   .689709452   .261433778
----------------------------+-----------------------------------
Treatment |        support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         31 |        31
   Treated |        81          3 |        84
-----------+----------------------+----------
     Total |        81         34 |       115

Bootstrap statistics                              Number of obs    =       115
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    28  .2614338 -.2188049  .6989004  -1.172591   1.695459   (N)
------------------------------------------------------------------------------
```

Matching estimates with Mahalanobis metrics, without caliper, # of match=1, in log for large firms in the high-technology industry

```
----------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------
        logemp_d  Unmatched | .613762203   .256483274   .357278929
                        ATT | .613762203   .274745477   .339016726
----------------------------+-----------------------------------
Treatment |  support
assignment | On suppor |     Total
-----------+-----------+----------
 Untreated |        86 |        86
   Treated |       225 |       225
-----------+-----------+----------
     Total |       311 |       311

Bootstrap statistics                              Number of obs    =      1502
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .3390167  .0270371  .1490606   .0394682   .6385653   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=1,in log for large firms in the high-technology industry

```
Probit estimates                                  Number of obs   =        311
                                                  LR chi2(16)     =     209.12
                                                  Prob > chi2     =     0.0000
Log likelihood = -78.818508                       Pseudo R2       =     0.5702
------------------------------------------------------------------------------
         tst |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lagemp |   .0443093   .0311334     1.42   0.155     -.016711    .1053297
       lagrd |   .0000444   .0014689     0.03   0.976    -.0028346    .0029235
    lagsales |  -.0001243   .0001136    -1.09   0.274     -.000347    .0000984
  knowcounty |    .000328   .0001194     2.75   0.006      .000094     .000562
   emp2digit |   2.58e-07   1.39e-06     0.19   0.853    -2.47e-06    2.99e-06
 firmsizeall |  -.0365133   .0305031    -1.20   0.231    -.0962982    .0232716
```

```
incapita89~n |  -.0000719   .0001016    -0.71   0.479     -.000271    .0001271
     edu1290 |   .0282586   .0429387     0.66   0.510    -.0558996    .1124168
     edu1690 |  -.0818808   .0413252    -1.98   0.048    -.1628766    -.000885
    sbir8690 |   8.70e-08   1.23e-08     7.07   0.000     6.29e-08    1.11e-07
    univ8690 |  -3.62e-06   6.52e-07    -5.55   0.000    -4.90e-06   -2.34e-06
   popchange |  -2.266128   .7673651    -2.95   0.003    -3.770136   -.7621199
     crime91 |  -.0000603   .0000616    -0.98   0.328    -.0001811    .0000605
   poverty89 |  -.0852375   .0691714    -1.23   0.218    -.2208109    .0503359
  unemprat90 |   -.256056   .1674963    -1.53   0.126    -.5843426    .0722307
mestdensi~87 |   321.7771   217.7516     1.48   0.139    -105.0082    748.5624
       _cons |   3.664107   3.980253     0.92   0.357    -4.137045    11.46526
------------------------------------------------------------------------------
```

```
----------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------------
        logemp_d  Unmatched | .613762203   .256483274   .357278929
                        ATT | .379368395   .329870647   .049497748
----------------------------+-----------------------------------------
 Treatment |         support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         86 |        86
   Treated |       192         33 |       225
-----------+----------------------+----------
     Total |       192        119 |       311

Bootstrap statistics                              Number of obs    =       311
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0494977  .0799642  .3969492  -.7482016   .8471971   (N)
------------------------------------------------------------------------------
```

Matching estimates with propensity score metrics, with caliper, # of match=3,in log for large firms in the high-technology industry

```
----------------------------------------------------------------------
        Variable     Sample |    Treated     Controls   Difference
----------------------------+-----------------------------------------
        logemp_d  Unmatched | .613762203   .256483274   .357278929
                        ATT | .379368395   .326125125   .053243271
----------------------------+-----------------------------------------
 Treatment |         support
assignment | Off suppo  On suppor |     Total
-----------+----------------------+----------
 Untreated |         0         86 |        86
   Treated |       192         33 |       225
-----------+----------------------+----------
     Total |       192        119 |       311

Bootstrap statistics                              Number of obs    =       311
                                                  Replications     =        50
------------------------------------------------------------------------------
Variable     |  Reps  Observed      Bias  Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    50  .0532433  .3201859  .3623146  -.6748551   .7813416   (N)
------------------------------------------------------------------------------
```

BIBLIOGRAPHY

Abadie, A. and Imbens, G. (2002) “Simple and bias-corrected matching estimators for average treatment effects,” *National Bureau of Economic Research*, Technical Working Paper No. 283, Cambridge, MA.

Abdel-Rahman, H. and Fujita, M. (1990) “Product variety, marshallian externalities, and city sizes,” *Journal of Regional Science*, Vol. 30, pp. 165-183.

Acs, Z. and Audretsch, D. (1987) “Innovation, market structure, and firm size,” *The Review of Economics and Statistics*, Vol. 69, pp. 567-574.

Acs, Z. and Audretsch, D. (1990) *Innovation and small firms*, Cambridge, MA, MIT Press.

Acs, Z. and Audretsch, D. (1991) “R&D, firm size and innovative activity", in Acs, Z. and Audretsch, D. (eds.) *Innovation and technological change: An international comparison*, Ann Arbor, University of Michigan Press.

Acs, Z., Audretsch, D., and Feldman, M. (1994) “R&D spillover and recipient firm size,” *The Review of Economics and Statistics*, Vol. 76, pp 336-40.

Aghion, P. and Howitt, P. (1992) “A model of growth through creative destruction,” *Econometrica*, Vol. 60, pp. 323-51.

Almeida, P. and Kogut, B. (1994) “Technology and geography: The localization of knowledge and the mobility of patent holders,” Working Paper, The Wharton School of Business.

Almeida, P. and Kogut, B. (1999) “Localization of knowledge and the mobility of engineers in regional networks,” *Management Science*, Vol. 45, pp. 905-917.

Altshuler, R. (1989) “A dynamic analysis of the research and experimentation credit,” *National Tax Journal*, Vol. 41, pp. 453-466.

Aman, G., Downey, G., and Panek, S. (2005) “Comprehensive revision of Gross State Product: Accelerated estimates for 2003, revised estimates for 1977-2002,” *Survey of Current Business*, Vol. 85, pp. 80 -106.

Angrist, J. and Hahn, J. (1999) "When to control for covariates? Panel-asymptotic results for estimates of treatment effect," *National Bureau of Economic Research*, Technical Working Paper No. 241, Cambridge, MA.

Angrist, J. and Krueger, A. (1998) "Empirical strategies in labor economics," in Ashenfelter, O. and Card, D. (eds.), *Handbook of labor economics*, North Holland, Amsterdam.

Anselin, L., Varga, A., and Acs, Z. (1997) "Local geographic spillovers between university research and high technology innovations," *Journal of Urban Economics*, Vol.42, pp. 422-448.

Antonelli, C. (1989) "A failure-inducement model of research and development expenditure," Journal of Economic Behavior and Organization, Vol. 12, pp. 159-180.

Arnold, L. (1998) "Growth, welfare, and trade in an integrated model of human capital accumulation and R&D," Journal of Macroeconomics, Vol. 20, pp.81-105.

Arrow, K. (1962a) "The economic implications of learning by doing," Review of Economic Studies, Vol. 29, pp.155-73.

Arrow, K. (1962b) "Economic welfare and the allocation of resources to invention," in Nelson, R. (eds.) *The rate and direction of inventive activity*, Princeton Univ. Press, pp. 609-625.

Audretsch, D. (1991) "New-firm survival and the technological regime," *The Review of Economics and Statistics*, Vol. 73, pp. 441-450.

Audretsch, D. (1995) "Innovation, growth and survival," *International Journal of Industrial Organization*, Vol. 13, pp. 441-45.

Audretsch, D. and Feldman, M. (1996) "R&D spillovers and the geography of innovation and production," *American Economic Review*, Vol. 86, pp. 630-640.

Baily, M. and Lawrence, R. (1987) "Tax policies for innovation and competitiveness," Study commissioned by the Council on Research and Technology, Washington, DC.

Baily, M. and Lawrence, R. (1992) "Tax incentives for R&D: What do the data tell us?," Study commissioned by the Council on Research and Technology, Washington, DC.

Barro, R. and Sala-i-Martin, X. (1991) "Convergence across states and regions," Brookings Papers on Economic Activity, pp. 107-82.

Bartik, T. (1984) "Business locational decisions in the U.S.: Estimates of the effects of unionization, taxes, and other characteristics of states," *Journal of Business and Economic Statistics*, Vol. 3, pp. 14-22.

Bartik, T. (1991) *Who benefits from state and local economic development policies*, Kalamazoo, Michigan, W.E. Upjohn Institute for Employment Research.

Bartik, T. (2002) "Evaluating the impacts of local economic development policies on local economic outcomes: What has been done and what is doable?," Upjohn Institute Staff Working Paper No. 03-89.

Bartik, T. and Bingham, R. (1997) "Can economic development programs be evaluated?," in Bingham R. and Mier, R. (eds) *Dilemmas of urban economic development: Issues in theory and practice*, Thousand Oaks, California, Sage Publications.

Bartelsman, E. (1990) "Federally sponsored R&D and productivity growth," Federal Reserve Economics Discussion Paper No. 121. Federal Reserve Board of Governors, Washington, DC.

Becker, S. and Ichino, A. (2002) "Estimation of average treatment effects based on propensity score," *The Stata Journal*, Vol. 2, pp. 358-377.

Beeson, P., Dejong, D., and Troesken, W. (2001) "Population growth in U.S. counties, 1840-1990," *Regional Science and Urban Economics*, Vol. 31, pp. 669-699.

Beeson, P. and Eberts, R. (1989) "Identifying productivity and amenity effects in interurban wage differentials," *The Review of Economics and Statistics*, Vol. 71, pp. 443-52.

Benus, J., Grover, N., Johnson, T., Shen, T. and Wood, M. (1995) "Self-Employment Programs: A new reemployment strategy, Final report on the UI Self-Employment Demonstration," the U.S. Department of Labor, Employment and Training Administration, Unemployment Insurance Service, Occasional Paper 95-4, Washington D.C., accessible via Internet WWW page, at URL: http://wdr.doleta.gov/research/FullText_Documents/op%5F04%2D95%2Epdf

Berger, P. (1993) "Explicit and implicit effects of the R&D tax credit," *Journal of Accounting Research*, Vol. 31, pp. 131-171.

Bernstein, J. and Nadiri, I. (1988) "Interindustry R&D spillovers, rates of return, and production in high-tech industries,' *American Economic Review*, Vol. 78, pp 429-34.

Bertrand, M., Duflo, E., and Mullainathan S. (2002) "How much should we trust differences-in-differences estimates?," *National Bureau of Economic Research*, Working Paper No. 8841, Cambridge, MA.

Bingham, R. and Mier, R. (eds.) (1993) *Theories of local economic development: Perspectives from across the disciplines*, Newbury Park, CA, Sage Publications.

Blackburn, K., Hung V., and Pozzolo, A. (2000) "Research, development and human capital accumulation," *Journal of Macroeconomics*, Vol. 22, pp.189-206.

Blakely, E. (1994) *Planning local economic development: Theory and practice, 2nd edition,* Thousand Oaks, California, Sage Publications.

Blakely, E. and Bradshaw, T. (2002) *Planning local economic development: Theory and practice, 3rd edition,* Thousand Oaks, California, Sage Publications.

Blair, J. and Premus, R. (1987) "Major factors in industrial location: A review," *Economic Development Quarterly*, Vol. 1. pp. 72-85.

Blair, J. and Premus, R. (1993) "Location Theory," in Bingham, R. and Mier, R. (eds.) *Theories of local economic development: Perspectives from across the disciplines*, London, Sage Publications, pp. 3-26.

Blanchard, O. and Katz, L. (1992) "Regional evolutions," *Brookings Papers on Economic Activity*, Vol. 1, pp. 1-61.

Bloom, N., Griffith, R., and Van Reenen, J. (2002) "Do R&D tax credits work? Evidence from a panel of countries 1979–1997," *Journal of Public Economics*, Vol. 85, pp. 1-31.

Blundell, R. and Costa Dias, M. (2000) "Evaluation methods for non-experimental data," *Fiscal Studies*, Vol. 21, pp. 427-468.

Bondonio, D. and Engberg, J. (2000) "Enterprise zones and local employment: Evidence from the states' programs," *Regional Science and Urban Economics*, Vol. 30, pp. 519-549.

Bound, J., Cummings, C., Griliches, Z., Hall, B., and Jaffe, A. (1984) "Who does R&D and who patents?" in Griliches, Z. (eds.) *R&D, patents, and productivity*, Chicago, University of Chicago Press.

Bozeman, B. and Link, A. (1984) "Tax incentives for R&D: A critical evaluation," *Research Policy*, Vol. 13, pp. 21-31.

Branstetter, L. and Sakakibara, M. (1998) "Japanese research consortia: A microeconometric analysis of industrial policy," *Journal of Industrial Economics*, Vol. 46, pp. 207-233.

Brown, S., Kathy, J., and Taylor, L. (2002) "State and local policy, factor markets and regional growth," Federal Reserve Bank of Dallas.

Browne, L. (1983) "High technology and business services," *New England Economic Review*, July/August, pp. 5-17.

Bucci, A. (2002) "When Romer meets Lucas: On human capital, imperfect competition and growth," Departemental Working Papers 2002-06, Department of Economics University of Milan Italy.

Busom, I. (2000) "An empirical evaluation of the effects of R&D subsidies," *Economics of Innovation and New Technology*, Vol. 9, pp. 111-148.

Bureau of Economic Analysis (2005) "State personal income 2005," News Release, BEA 06-10, Project Officer, D. Lenze and K. Albetski.

Buxton, A. (1975) “The process of technical change in UK manufacturing,” *Applied Economics*, Vol. 7, pp. 53-71.

Caballero, R. and Jaffe, A. (1993) “How high are the giant’s shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth,” NBER Macroeconomic Annual, Cambridge, MIT Press.

Campbell, D. (1957) “Factors relevant to the validity of experiments in social settings,” *Psychological Bulletin*, Vol. 54, pp. 297-312.

Campbell, D. and Stanley, J. (1963) *Experimental and quasi-experimental designs for research,* Chicago, Rand McNally College Publishing Company.

California Council on Science and Technology (1999), *California Report on the Environment for Science and Technology,* accessible via Internet WWW page, at URL: http://www.ccst.us/ccst/pubs/crest/pubs/creports/CREST_Re.pdf

Card, D. (1990) “The impact of the Mariel Boatlift on the Miami Labor Market,” *Industrial and Labor Relations Review*, Vol. 43, pp. 245-57.

Card, D. and Krueger, A. (1994) “Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania,” *American Economic Review*, Vol. 84, pp. 772-793.

Carlton, D. (1983) “The location and employment choices of new firms: An econometric model with discrete and continuous endogenous variables,” *The Review of Economics and Statistics*, Vol. 65, pp. 440-449.

CCH, Tax Research Network, accessible via Internet WWW page, at URL: http://tax.cch.com/primesrc/bin/highwire.dll?ult=p&tpl=drilogin.tpl&>

Chakrabarti, A. (1991) “Industry characteristics influencing the technical output: A case of small and medium-sized firms in the U.S,” *R&D Management*, Vol. 21, pp. 139-152.

Charlot, S. and Duranton, G. (2004) “Communication externalities in cities,” *Journal of Urban Economics*, Vol. 56, pp. 581-613.

Chinitz, B. (1961) “Contrasts in agglomeration: New York and Pittsburgh,” *American Economic Review*, Vol. 51, pp. 279-289.

Christaller, W. (1933) *Central places in southern Germany*, Translated (in part) by Charlisle B. (1966), Englewood Cliffs, Prentice-Hall.

Ciccone, C. and Hall, R. (1996) “Productivity and the density of economic activity,” *American Economic Review*, Vol. 86, pp. 54-70.

Collins, W. (2003) “The labor market impacts of state-level anti-discrimination laws,” *Industrial and Labor Relations Review*, Vol. 56, pp. 244-272.

Cohen, W. and Levinthal, D. (1989) "Innovation and learning: The two faces of R&D," *Economic Journal*, Vol. 99, pp. 569-596.

Cohen, W. and Levin, R. (1989) "Empirical studies of innovation and market structure," in Schmalensee, R. and Willig, R. (eds.) *Handbook of industrial organization,* New York, North-Holland.

Cohen, W., Levin, R., and Mowery, D. (1987) "Firm size and R&D intensity: A re-examination," *Journal of Industiral Economics*, Vol. 35, pp. 543-563.

Cook, T. and Campbell, D. (1979) *Quasi-experimentation: Design & analysis issues for field settings*, Boston, Houghton Mifflin Company.

Cordes, J. (1989) "Tax incentives and R&D spending: A review of the evidence," *Research Policy*, Vol. 18, pp. 119-133.

Crosby, M. (2000) "Patents, innovation and growth," *Economic Record*, Vol. 76, pp. 255-262.

David, P., Hall, B., and Toole, A. (2000) "Is public R&D a complement or substitute for private R&D? A review of the econometric evidence," *Research Policy*, Vol. 29, pp.497-529.

Dehejia, R. and Wahba, S. (1999) "Causal effects in nonexperimental studies: Reevaluation of the evaluation of training programs," *Journal of the American Statistical Association*, Vol. 94, pp.1053-1062.

Dehejia, R. and Wahba, S. (2002) "Propensity score matching method for nonexperimental causal studies," *The Review of Economics and Statistics*, Vol. 84, pp.151-161.

Denison, E. (1967) *Why growth rates differ: Postwar experience in nine western counties*, Washington, The Brookings Institution.

Department of Revenue of Washington State (1997) "High technology R&D tax incentives study;" the Department of Revenue, the Research Division, accessible via Internet, at URL:http://dor.wa.gov/content/statistics/1997/High_Tech_RandD_Study_1997/default.aspx

Department of Revenue of Washington State (2000) "High technology R&D tax incentives study;" the Department of Revenue, the Research Division, accessible via Internet, at URL:http://dor.wa.gov/content/statistics/2000/High_Tech_RandD_Study_2000/default.aspx

Department of Revenue of Washington State (2003) "High technology R&D tax incentives study;" the Department of Revenue, the Research Division, accessible via Internet, at URL:http://dor.wa.gov/content/statistics/2003/High_Tech_RandD_Study_2003/default.aspx

DeVol, R. (1999) *America's high-tech economy: Growth, development, and risks for metropolitan areas*, Santa Monica, The Milken Institute.

Diamond, A. (1998) "Does federal funding crowd out private funding of science?," Presented at the American Economics Association meetings, Chicago, January.

Dowall, D. "An evaluation of California's enterprise zone programs," *Economic Development Quaterly*, Vol. 10, pp. 352-268.

Drucker, P. (1991) "The changed world economy," in Fosler, S. (eds.) *Local economic development*, Washington D.C., International City Management Association.

Dumais, G., Ellison, G., and Glaeser, E. (2002) "Geographic concentration as a dynamic process," *The Review of Economics and Statistics*, Vol. 84, pp. 193-204.

Duranton, G. and Puga, D. (2000) "Diversity and specialization in cities: Why, where and when does it matter?," *Urban Studies*, Vol. 37, pp. 533-555.

Duranton, G. and Puga, D. (2001) "Nursery cities: Urban diversity, process innovation, and the life cycle of products," *American Economic Review*, Vol. 91, pp. 1454-1477.

Eaton, J. and Eckstein, Z. (1997) "Cities and growth: Theory and evidence from France and Japan," *Regional Science and Urban Economics*, Vol. 27, pp 443-474.

Eisenger, P. (1995) "State economic development in the 1990s: Politics and policy learning," *Economic Development Quarterly*, Vol. 9, pp. 146-158.

Eisner, R., Steven, H., and Sullivan, M. (1983) "Tax incentives and R&D expenditures", in Ecole Nationale de la Statistique et de l'Administration Economique and National Bureau of Economic Research, *Proceedings of the Conference on Quantitative Studies of Research and Development in Industry*, Vol. 2. CNRS, Paris, France, pp. 375-466.

Eisner, R., Albert, S., and Sullivan, M. (1984) "The new incremental tax credit for R&D: Incentive or disincentive?" *National Tax Journal*, Vol. 37, pp. 171-185.

Elling, R. and Sheldon, A. (1991) "Comparative analyses of state enterprise zone programs," in Green, R. (eds.) *Enterprise zones,* Newbury Park, Sage Publications, pp. 136-154.

Ellison, G. and Glaeser, E. (1997) "Geographic concentration in U.S. manufacturing industries: A dartboard approach," *Journal of Political Economy*, Vol. 105, pp. 889-927.

Feldman, M. (1994) *The Geography of Innovation*, London, Kluwer Academic.

Feldman, M. and Audretsch, D. (1999) "Innovation in cities: Science-based diversity, specialization and localized competition," *European Economic Review*, Vol. 43, pp. 409-429.

Fernald, J. (1999) “Roads to prosperity: Assessing the link between public capital and productivity,” *American Economic Review*, Vol. 89, pp. 619-38.

Friedlander, D. and Robins, P. (1995) “Evaluating program evaluations: New evidence on commonly used nonexperimental methods,” *American Economic Review*, Vol. 85, pp. 923-937.

Gabaix, X (1999a) “Zipf’s law for cities: An explanation,” *The Quarterly Journal of Economics*, Vol. 114, pp. 739-767.

Gabaix, X. (1999b) “Zipf's law and the growth of cities,” *American Economic Review*, Vol. 89, pp. 129-132.

Galbraith, J. (1952) *American capitalism, the concept of countervailing power*, Boston, Houghton Mifflin Company.

Garcia-Milà, T., McGuire, T., and Porter, R. (1996) “The effect of public capital in state-level production functions reconsidered,” *The Review of Economics and Statistics*, Vol. 78, pp. 177-180.

Garcia-Milà, T. and McGuire, T. (2001) “Tax incentives and the city”, Economic Working Papers No. 631, Department of Economics and Business, Universitat Pompeu Fabra.

Gaspar, J. and Glaeser, E. (1998) “Information technology and the future of cities,” *Journal of Urban Economics*, Vol. 43, pp. 136-156.

Geroski, P., Machin, S., and Van Reenen, J. (1993) “Innovation and firm profitability,” *RAND Journal of Economics*, Vol. 24, pp. 198-211.

Geroski, P. (1995) “Do spillovers undermine the incentive to innovate?,” in Dowrick, S. (eds.) *Economic approaches to innovation*, Edwar Elgar, Aldershot, pp. 76-97.

Gibrat, R. (1931) *Les inegalities economiques*, Paris, Librairie du Recueil Sirey.

Glaeser, E. (1994) “Cities, information and economic growth,” *Citiscape*, Vol. 1 pp. 9-48.

Glaeser, E. (1999) “Learning in Cities,” *Journal of Urban Economics*, Vol. 46, pp. 254-277.

Glaeser, E. (2000) “The new economics of urban and regional growth,” in Clark, G., Feldman, M., and Gertler, M. (eds.) *The Oxford handbook of economic geography*, Oxford, England, The Oxford University Press, pp. 83-98.

Glaeser, E., Kallal, H., Scheinkman, J., and Shleifer, A. (1992) “Growth in cities,” *Journal of Political Economy*, Vol. 100, pp. 1126 -1152.

Glaeser, E., Scheinkman, J., and Shleifer, A. (1995) “Economic growth in a cross-section of cities,” *Journal of Monetary Economics*, Vol. 36, pp. 117-143.

Glasmeier, A. (1988) “Factors governing the development of high tech industry agglomerations: A tale of three cities,” *Regional Studies*, Vol. 22, pp.287-301.

Globerman, S. (1973) “Market structure and R&D in Canadian manufacturing industries,” *Quarterly Review of Economics and Business*, Vol. 13, pp. 59-68.

Goldstein, H. and Lugar, M. (1997) “Theory and practice in high-tech economic development,” in Bingham, R. and Mier R. (eds.) *Theories of local economic development: Perspectives from across the disciplines*, Newbury Park, CA, Sage publications.

Gray, W. (1997) “Manufacturing plant location: Does state pollution regulation matter?,” *NBER Working Paper* 5880.

Greenbaum, R. and Engberg, J. (1998) “The impact of state urban enterprise zones on business outcomes,” *Center for Economic Studies*, Discussion paper No. 98-20.

Griliches, Z. (1979) “Issues in assessing the contribution of R&D to productivity growth,” *The Bell Journal of Economics*, Spring, pp. 92-116.

Griliches, Z. and Lichtenberg, F. (1984) “R&D and productivity growth at the industry level: Is there still a relationship?,” in Griliches, Z. (Eds.) *R&D, patents and productivity,* Chicago, University of Chicago Press.

Griliches, Z. (1986) “Productivity, R&D and basic research at the firm level in the 1970s,” *American Economic Review*, Vol. 76, pp. 141-154.

Griliches, Z. (1992) “The search for R&D spillovers,” *Scandinavian Journal of Economics*, Vol. 94, pp. S29-S47.

Griliches, Z. (1995) “R&D and productivity: Econometric results and measurement issues,” in Paul, S. (eds.) *The handbook of the economics of innovation and technological change*, Oxford, Blackwell.

Griliches, Z. and Regev, H. (1998) “An econometric evaluation of high-tech policy in Israel,” Paper presented at ATP-conference in Washington, DC, June 1998.

Grossman, G. and Helpman, E. (1991a) *Innovation and growth in the global economy*, Cambridge, MA, MIT Press.

Grossman, G. and Helpman, E. (1991b) “Trade, knowledge spillovers, and growth,” *European Economic Review*, Vol. 35, pp. 517-26.

Gruber, J. (1994) “The incidence of mandated maternity benefits,” *American Economic Review*, Vol. 84, pp. 622-641.

Hadlock, P., Hecker, D., and Gannon, J. (1991) “High technology employment: Another view,” *Monthly Labor Review*, July, pp. 26-30.

Hall, B. (1993) "R&D tax policy during the 1980s: Success or failure," *Tax Policy and the Economy*, Vol. 7, pp. 1-35.

Hall, B. (1996) "The private and social returns to research and development," in Smith, B. and Barfield, C. (eds.) *Technology, R&D and the economy*, AEI-Brookings Institution, Washington, DC.

Hall, B., Jaffe, A., and Trajtenberg, M. (2001) "The NBER patent citation data file: Lessons, insights and methodological tools," NBER Working Paper, No. w8498.

Hall, B. and Van Reenen, J. (2000) "How effective are fiscal incentives for R&D? A review of the evidence," *Research Policy*, Vol. 29, pp. 449-469.

Hall, B. and Wosinska, M. (1999) "The California R&D tax credit: Description, history and economic analysis," California Council on Science and Technology Series.

Harrison, B., Kelley, M., and Gant, J. (1996) "Innovative firm behavior and local milieu: Exploring the intersection of agglomeration, firm effects, and technological change," *Economic geography*, Vol. 72, pp. 233-258.

Haughwout, A. (2002) "Public infrastructure investments, productivity and welfare in fixed geographic areas," *Journal of Public Economics*, Vol. 83, pp.405-428.

Hausman, J., Hall, B., and Griliches, Z. (1984) "Econometric models for count data with an application to the patents-R&D relationship," *Econometrica*, Vol. 52, pp.909-938.

Heckman, J., Ichimura, H., Smith, J., and Todd, P. (1998) "Characterizing selection bias using experimental data," *Econometrica,* Vol. 66, pp 1017-1098.

Heckman, J., Hidehiko, I., and Todd, P. (1997) "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme," *Review of Economic Studies,* Vol. 64, pp. 605-654.

Helsley, R. and Strange, W. (1990) "Matching and agglomeration economies in a system of cities," *Regional Science and Urban Economics*, Vol. 20, pp. 189-212.

Helsley, R. and Strange, W. (2004) "Knowledge barter in cities," *Journal of Urban Economics*, Vol. 56, pp. 327-345.

Henderson, V. (1986) "Efficiency of resource usage and city size," *Journal of Urban Economics*, Vol. 19, pp. 47-70.

Henderson, V. (1988) *Urban development: Theory, fact and illusion*, New York, Oxford University Press.

Henderson, V. (2003) "Marshall's scale economies," *Journal of Urban Economics*, Vol. 56, pp. 1-28.

Henderson, V., Kuncoro, A., and Turner, M. (1995) "Industrial development in cities," *Journal of Political Economy*, Vol. 103, pp 1067-1090.

Hines, J. (1993) "No place like home: Tax incentives and the location of R&D by American multinationals," *Tax Policy and the Economy*, Vol. 8, pp. 65-104.

Holmes, T. (1998) "The effect of state policies on the location of manufacturing: Evidence from state borders," *Journal of Political Economy*, Vol. 103, pp. 1067-1090.

Hoover, E. (1948) *The location of economic activity*, New York, McGraw-Hill.

Hoover, E. (1971) *An introduction to regional economic*, New York, Knopf.

Howe, J. and McFetridge, D. (1976) "The determinants of R&D expenditures," *Canadian Journal of Economics*, Vol. 9, pp. 57-71.

Irwin, D. and Klenow, P. (1996) "High-tech R&D subsidies – estimating the effects of SEMATECH," *Journal of International Economics*, Vol. 40, pp. 323-344.

Isard, W. (1956) *Location and space economy,* Cambridge, MA, MIT Press.

Isserman, A. and Merrifield, J. (1987) "Quasi-experimental control group methods for regional analysis: An application to an energy boomtown and growth pole theory," *Economic Geography*, Vol. 63, pp.3-19.

Isserman, A. and Rephann, T. (1995) "The economic effects of the Appalachian regional commission," *Journal of the American Planning Association*, Vol. 61, pp. 345-364.

Jacobs, J. (1969) *The economy of cities,* New York, Random House.

Jaffe, A. (1986) "Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits and market value," *American Economic Review*, Vol. 76, pp. 984-100.

Jaffe, A. (1989) "Real effects of academic research," *The American Economic Review*, Vol. 79, pp. 957-970.

Jaffe, A., Trajtenberg, M., and Henderson, R. (1993) "Geographic localization of knowledge spillovers as evidenced by patent citations," *Quarterly Journal of Economics*, Vol. 63, pp. 77-598.

Jones, L. and Manuelli, R. (1990) "A convex model of equilibrium growth," *Journal of Political Economy*, Vol. 98, pp. 1008-1038.

Jones, C. and Williams, J. (1998) "Measuring the social return to R&D," *Quarterly Journal of Economics*, Vol. 113, pp. 1119-35.

Kao, C. (1999) "Spurious regression and residuals based tests for cointegration in panel data," *Journal of Econometrics*, Vol. 90, pp. 1-44.

Klenow, P. (1998) "Ideas *vs.* rival human capital: Industry evidence on growth models," *Journal of Monetary Economics*, Vol. 42, pp. 3-24.

Klette, T. and Moen, J. (1998) "R&D investment responses to R&D subsidies: A theoretical analysis and econometric evidence," presented to the NBER Summer Institute, July.

Klette, T., Møen, J., and Griliches, Z. (2000), "Do subsidies to commercial R&D reduce market failures? Microeconomic evaluation studies," *Research Policy*, Vol. 29, pp 473-497.

Koga, T. (2003) "Firm size and R&D tax incentives," *Technovation*, Vol. 23, pp. 643-648.

Krugman, P. (1991a) *Geography and trade*, Cambridge, MA: MIT Press.

Krugman, P. (1991b) "Increasing returns and economic geography," *Journal of Political Economy*, Vol. 99, pp. 483-499.

LaLonde, R. (1986) "Evaluating the econometric evaluations of training programs with experimental data," *American Economic Review*, Vol. 76, pp. 604-20.

Lerner, J. (1999) "The government as venture capitalist: the long-run impact of the SBIR program," *Journal of Business*, Vol. 72, pp. 285-318.

Lee, C. and Sung, T. (2005) "Schumpeter's legacy: A new perspective on the relationship between firm size and R&D," *Research Policy*, Vol. 34, pp. 914-931.

Leuven, E. and Sianesi, B. (2003) "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing," accessible via Internet WWW page, at URL: http://ideas.repec.org/c/boc/bocode/s432001.html Version 1.2.3.

Levin, R. and Reiss, P. (1984) "Tests of a Schumpeterian model of R&D and market structure," in Griliches, Z. (eds.) *R&D, patents and productivity*, Chicago, University of Chicago Press.

Levy, D. (1990) "Estimating the impact of government R&D," *Economic Letters*, Vol. 32, pp. 169-173.

Levy, D. and Terleckyj, N. (1983) "Effects of government R&D on private R&D investment and productivity: A macroeconomic analysis," *Bell Journal of Economics*, Vol. 14, pp. 551-561.

Leyden, D. and Link, A. (1991) "Why are government and private R&D complements?," *Applied Economics*, Vol. 23, pp. 1673-1681.

Leyden, D. and Link, A. (1993) "Tax policies affecting R&D: An international comparison," *Technovation*, Vol. 13, pp.17-25.

Lichtenberg, F. (1984) “The relationship between federal contract R&D and company R&D,” *American Economic Review Papers and Proceedings*, Vol. 74, pp. 73-78.

Lichtenberg, F. (1987) “The effect of government funding on private industrial research and development: A re-assessment,” *The Journal of Industrial Economics*, Vol. 36, pp. 97-104.

Lichtenberg, F. (1988) “The private R&D investment response to federal design and technical competitions,” *American Economic Review*, Vol. 78, pp. 550-559.

Link, A. (1980) “Firm size and efficient entrepreneurial acitivity: A reformation of the schumpetre hypothesis,” *Journal of Political Economy*, Vol. 88, pp. 771-782.

Link, A. (1982) “An analysis of the composition of R&D spending,” *Southern Journal of Economics*, Vol. 49, pp. 342-349.

Lösch, A. (1954) *The Economics of Location (english version),* New Haven, Yale University Press.

Lucas, R. (1993) “Making a miracle,” *Econometrica*, Vol. 61, pp. 251-272.

Lucas, R. (1988) “On the mechanics of economic development,” *Journal of Monetary Economics*, Vol. 22, pp. 3-42.

Lucas, R. (2001) “Externalities and cities,” *Review of Economic Dynamics*, Vol. 4, pp. 245-274.

Lugar, M. and Goldstein, H. (1991) *Technology in the garden*, Chapel Hill, The University of North Carolina Press.

Luker, W. Jr. and Lyons, D. (1997) “Employment shifts in high-technology industries, 1988-1996,” *Monthly Labor Review,* June, pp. 12-25.

Malecki, E. (1984) “High technology and local economic development,” *Journal of the American Planning Association*, Vol. 50, pp. 260-69.

Malecki, E. (1997) *Technology and economic development: the dynamics of local, regional and national competitiveness*, 2nd edition, Essex, England: Addison Wesley Longman.

Malizia, E. and Feser, E. (1999) *Understanding local economic development,* New Brunswick, N.J., Center for Urban Policy Research, Rutgers University.

Mamuneas, T. and Nadiri, I. (1996) “Public R&D policies and cost behavior of the US manufacturing industries,” *Journal of Public Economics*, Vol. 63, pp. 57-81.

Mansfield, E. (1965) “Rates of return from industrial research and development,” *American Economic Review*, Vol. 55, pp. 310-22.

Mansfield, E. (1986) “The R&D tax credit and other technology policy issues,” *American Economic Association Papers and Proceedings*, Vol. 76, pp. 190-194.

Mansfield, E. and Switzer, L. (1985) "The effects of R&D tax credits and allowances in Canada," *Research Policy,* Vol.14, pp.97-107.

Mark, S., McGuire, T., and Papke, L. (2000) "The influence of taxes on employment and population growth: Evidence from the Washington, D.C. metropolitan area," *National Tax Journal*, Vol. 55, pp. 105-124.

Markusen, A. (1985) *Profit cycles, oligopoly, and regional development*, Cambridge, MA, MIT Press.

Markusen, A. (1996) "Sticky places in slippery space: A typology of industrial districts," *Economic Geography*, Vol. 72, pp. 293-313.

Markusen, A., Hall, P., and Glasmeier, A. (1986) *High tech America: The what, how, where and why of the sunrise industries*, Boston, Allen and Unwin.

Marshall, A. (1890) *Principles of economics: An introductory volume,* 1990 reprint of 1920 edition, Philadelphia, Porcupine.

Massachusetts Technology Collaborative (2002) *Maintaining the innovation edge*,

Massachusetts Technology Collaborative (2004) *The R&D funding scorecard: Federal investments and the Massachusetts innovation economy,*

McCain, L. and McCleary, R. (1979) "The statistical analysis of the simple interrupted time-series quasi-experiment," in Cook, T. and Campbell, D., *Quasi-experimentation: Design & analysis issues for field settings*, Boston, Houghton Mifflin Company.

McCutchen, W. (1993) "Estimating the impact of the R&D tax credit on strategic groups in the pharmaceutical industry," *Research Policy*, Vol. 22, pp. 337-351.

Melkers, J., Bugler, D., and Bozeman, B. (1997) "Technology transfer and economic development," in Bingham, R. and Mier, R. (eds.) *Theories of local economic development: Perspectives from across the disciplines*, Newbury Park, California, Sage Publications.

Meyer, B. (1995) "Natural and quasi-experiments in economics," *Journal of Business and Economic Statistics*, Vol. 13, pp. 151-161.

Meyer, B. and Sullivan, J. (2004) "The effects of welfare and tax reform: The material well-being of single mothers in 1980s and 1990s," *Journal of Public Economics*, Vol. 88, pp. 1387-1420.

Michigan Economic Development Corporation (2002) *Next Michigan: Our competitive action agenda*, accessible via Internet WWW page: at URL: http://medc.michigan.org/news/reports/economic/combo.asp?ContentId=066B0AE4-A6E4-4014-B290-57BFF6ABB91A&QueueId=1&ContentTypeId=9

Mody, A. and Wang, F. (1997) "Explaining industrial growth in coastal China: Economic reforms…and what else?," *World Bank Economic Review*, Vol. 11, pp.293-325.

Mokyr, J. (1990) *The lever of riches: Technological creativity and economic progress*, New York, Oxford University Press.

Mowery, D. and Rosenberg, N. (1989) *Technology and the pursuit of economic growth*, New York, Cambridge University Press.

Mowery, D. and Ziedonis, A. (2001) "The geographic reach of market and non-market channels of technology transfer: Comparing citations and licenses of university patents," *National Bureau of Economic Research*, Working Paper No. 8568, Cambridge, MA.

Mueser, P., Troske, K., and Gorislavsky, A. (2003) "Using state administrative data to measure program performance," *Institute for the Study of Labor*, Discussion Paper No. 786.

Myers, D. (1987a) "Community–relevant measurement of quality of life: A focus on local trends," *Urban Affairs Quarterly*, Vol. 23, pp. 108-125.

Myers, D. (1987b) "Internal monitoring of quality of life for local economic development," *Economic Development Quarterly*, Vol. 1, pp. 268-278.

Nadiri, I. (1993) "Innovation and technological spillovers," *National Bureau of Economic Research,* Working Paper, No. 4423, Cambridge, MA.

Nadiri, I. and Mamuneas, T. (1994) "The effects of public infrastructure and R&D capital on the cost structure and performance of U.S. manufacturing industries," *Review of Economics and Statistics*, Vol. 76, pp. 22-37.

National Science Foundation, National Science Board (1993) *Science and Engineering Indicators 1993*, Arlington, VA.

National Science Foundation, Division of Science Resources Statistics (1995) "Six states account for majority of R&D spending, new NSF state science and engineering profiles available," Data Brief, NSF 95-338, Project Officer, Richard Bennof and Steven Payson, Arlington, VA.

National Science Foundation, Division of Science Resources Statistics (1998) "Six states account for half the nation's R&D," Data Brief, NSF 98-306, Project Officer, Richard Bennof and Steven Payson, Arlington, VA.

National Science Foundation, Division of Science Resource Studies (1999) "How has the field mix of federal research funding changed over the past three decades?," Issue brief, NSF 99-328, Project Officer, Alan I. Rapoport, Arlington, VA.

National Science Foundation, National Science Board (2000) *Science and engineering indicator 2000*, Arlington, VA.

National Science Foundation, Division of Science Resources Statistics (2001) “R&D spending is highly concentrated in a small number of states,” Data Brief, NSF 01-320, Project Officer, Richard Bennof, Arlington, VA.

National Science Foundation, Division of Science Resources Statistics (2002) *Research and development in industry: 1999,* NSF 02-312, Project Officer and Principal Author, Raymond M. Wolfe, Arlington, VA.

National Science Foundation, Division of Science Resources Statistics (2002) “Top R&D performing states display diverse R&D patterns in 2000,” Infobrief, NSF 03-030, Project Officer, Brandon Shackelfold, Arlington, VA.

National Science Foundation, Division of Science Resources Statistics (2004) *Federal Funds for Research and Development: Fiscal Years 2001, 2002,and 2003,* NSF 04-310, Project Officer, Ronald L. Meeks, Arlington, VA.

National Science Foundation, Division of Science Resources Statistics (2005) *National Patterns of Research and Development Resources: 2003,* NSF 05-308, Project Officer, Brandon Shackelford, Arlington, VA.

Nelson, R. (2000) “Research on productivity growth and productivity differences: dead ends and new departures,” in Nelson, R., (2000) *The sources of economic growth*, Cambridge, Harvard University Press.

Nitsch, V. (2005) “Zipf zipped,” *Journal of Urban Economics*, Vol. 56, pp. 86-100.

Office of Technology Assessment (1995a) *The effectiveness of research and experimental tax credits*, Washington, DC, U.S. Government Printing Office.

Office of Technology Assessment (1995b) *A history of the department of defense federally funded research and development centers,* OTA-BP-ISS-157, Washington, DC, U.S. Government Printing Office.

Ottaviano, G. and Puga, D. (1998) “Agglomeration in the global economy: A survey of the “New Economic Geography,” *World Economy*, Vol. 21, pp. 707-731.

Papke, L. (1991) “Interstate business tax differentials and new firm location: Evidence from panel data,” *Journal of Public Economics*, Vol. 45, pp. 47-68.

Papke, L. (1993) “What do we know about enterprise zones?,” in Poterba, J. (eds.) *Tax Policy and the Economy 7,* Cambridge, MIT Press, pp. 37-72.

Pavitt, K. (1983) “Characteristics of innovative activities in British industries,” *Omega*, Vol. 11, pp.113-130.

Phillips, P. and Moon, H. (1999) “Linear regression limit theory for nonstationary panel data,” *Econometrica*, Vol. 67, pp. 1057-1112.

Phillips, P. and Moon, H. (2000) “Nonstationary panel data analysis: an overview of some recent developments,” *Econometric Reviews*, Vol. 19, pp. 263-286.

Piore, M. and Sabel, C. (1984) *The second industrial divide*, New York, NY, Basic Books.

Porter, M. (1990) *The competitive advantage of nations*, New York, NY, The Free Press.

Porter, M. (2000) “Location, competition and economic development: Local clusters in a global economy,” *Economic Development Quarterly*, Vol. 14, pp. 15-34.

Rappaport, J. (1999) “Local growth empirics,” Center of International Development at Harvard University Working Paper No.23.

Rashkin, M. (2003) *Research and development tax incentives - federal, state, and foreign,* CCH Inc., Chicago, Wolters Kluwer Company.

Rauch, J. (1993) “Productivity gains from geographic concentration of human capital: Evidence from the cities,” *Journal of Urban Economics*, Vol. 34, pp. 380-400.

Rauch, J. (1995) “Bureaucracy, infrastructure, and economic growth: Evidence from US cities during the progressive era,” *American Economic Review*, Vol. 85, pp. 968-79.

Rebelo, S. (1991) “Long-run policy analysis and long-run growth,” *Journal of Political Economy*, Vol. 99, pp. 500-521.

Reed, R. and Rogers, C. (2003) “A study of quasi-experimental control group methods for estimating policy impacts,” *Regional Science and Urban Economics*, Vol. 33, pp 3-25.

Reese, L. and Fastenfest, D. (1999) “What works best?,” in Blair, J. and Reese, L. (eds.) *Approaches to economic development: Readings from economic development quarterly* , Thousand Oaks, California, Sage Publications.

Reichardt, C. (1979) “The statistical analysis of data from nonequivalent group designs,” in Cook, T. and Campbell, D., *Quasi-experimentation: Design & analysis issues for field settings*, Boston, Houghton Mifflin Company.

Rephann, T. and Isserman, A. (1994) “New highways as economic development tools: An evaluation using quasi-experimental matching methods,” *Regional Science and Urban Economics*, Vol. 24, pp. 723-751.

Riche, R., Hecker, D. and Burgan, J. (1983) “High technology today and tomorrow: A small slice of the employment pie,” *Monthly Labor Review*, pp. 50-58.

Romer, P. (1986) “Increasing returns and long run growth,” *Journal of Political Economy*, Vol. 94, pp. 1002-37.

Romer, P. (1990) “Endogenous technological change,” *Journal of Political Economy*, Vol. 98, pp. S71-S102.

Romer, P. (1993) "Idea gaps and object gaps in economic development," *Journal of Monetary Economics*, Vol. 32, pp. 543-574.

Rosenbaum, P. and Rubin, D. (1983) "The central role of the propensity score in observational studies for causal effects," *Biometrika*, Vol. 70, pp.41-55.

Rosenbaum, P. and Rubin, D. (1985) "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score," *The American Statistician*, Vol. 39, pp.33-38.

Rosenthal, S. and Strange, W. (2001) "The determinants of agglomeration," *Journal of Urban Economics*, Vol. 50, pp. 191-229.

Rubin, D. (1977) "Assignment to a treatment group on the basis of a covariate," *Journal of Educational Statistics*, Vol. 2, pp. 1-26.

Saiz, M. (2001) "Using program attributes to measure and evaluation state economic development strategies," *Economic Development Quarterly*, Vol. 15, pp. 45-57.

Scherer, F. (1983) "The propensity to patent," *International Journal of Industrial Organization*, Vol. 1, pp.107-128.

Scherer, F. (1984) *Innovation and growth: Schumpeterian perspectives*, Cambridge, The MIT Press.

Scherer, F. (1992) "Schumpeter and plausible Capitalism," *Journal of Economic Literature,* Vol. 30, pp. 1416-1433.

Schmenner, R. (1981) "Location decisions of large firms: Implications for public policy," *Commentary*, pp. 3-7.

Schmenner, R. (1982) *Making business locational decision*, Eaglewood Cliffs, NJ, Prentice-Hall.

Schumpeter, J. (1934) *The theory of economic development*, Cambridge, Harvard University Press, (New York: Oxford University Press, 1961.) First published in German, 1912.

Schumpeter, J. (1942) *Capitalism, socialism and democracy*, New York, Harper and Brothers. (Harper Colophon edition, 1976.)

Scott, A. (1988) *Metropolis: From the division of labor to urban form*, Berkeley, University of California Press.

Scott, J. (1984) "Firm versus industry variability in R&D intensity," in Griliches, Z. (eds.) *R&D, patents and productivity*, Chicago, University of Chicago Press.

Secretaries of Technology and Commerce and Trade of Virginia State (2000), "A joint study of research and development tax incentives," House Document No.46, Commonwealth of Virginia, Richmond.

Shadish, W., Cook, T., and Campbell, D. (2002) *Experimental and quasi-experimental designs for generalized causal inference*, Boston, Houghton Mifflin Company.

Siegel, P., Johnson, T., and Alwang, J. (1995) "'Regional economic diversity and diversication," *Growth and Change*, Vol. 24, pp. 261-84.

Sigalla, F. and Viard A. (1999) "Would a research tax credit be a good investment for Texas?," *Southwest Economy*, March/April, Federal Reserve Bank of Dallas.

Simon, C. (1998) "Human capital and metropolitan employment growth," *Journal of Urban Economics*, Vol. 43, pp. 223-243.

Simon, C. and Nardinelli, C. (1996) "The talk of the town: Human capital, information, and the growth of English cities, 1861 to 1961," *Explorations in Economic History*, Vol. 33, pp. 384-413.

Simon, C. and Nardinelli, C. (2002) "Human capital and the rise of American cities, 1900-1990," *Regional Science and Urban Economics*, Vol. 32, pp. 59-96.

Smith, P. (1999) "Do knowledge spillovers contribute to U.S. state output and growth?," *Journal of Urban Economics*, Vol. 45, pp. 331-353.

Solow, R. (1957) "Technical change and the aggregate production function," *The Review of Economics and Statistics*, Vol. 39, pp. 312-320.

Sorenson, D. (1997) "An empirical evaluation of profit cycle theory," *Journal of Regional Science*, Vol. 37, pp. 275-305.

Spence, M. (1984) "Cost-reduction, competition and industry performance," *Econometrica*, Vol. 52, pp.101-121.

State Science and Technology Institute (1997) "State research and development tax incentives," accessible via Internet WWW page, at URL: <http://www.ssti.org/Publications/Onlinepubs/r&d_97.pdf>

Stiglitz, J. (1988) *Economics of the Public Sector*, 2nd edition, New York, Norton and Company.

Stoneman, P. (1987) "Some aspects of the relation between technological change and economic performance," in *Economic analysis of technological policy*, Oxford, Clarendon Press.

Storper, M. and Walker, R. (1989) The capitalist imperative: Territory, technology, and industrial growth, New York, Basil Blackwell.

Sullivan, A. (1993) *Urban economics*, 2nd edition, Homewood, Illinois, Irwin.

Swenson, C. (1992) "Some tests of the incentive effects of the research and experimentation tax credit," *Journal of Public Economics*, Vol. 49, pp. 203-218.

Taylor, M. (1986) "The product-cycle model: A critique," *Environment and Planning A*, Vol. 18, pp. 751-761.

Taylor, M. (1987) "Enterprise and the product-cycle model: Conceptual ambiguities," in Van Der Knaap, B. and Wever, E. (eds.) *New technology and regional development*, London, Croom Helm, pp. 75-93.

Thisse, J. (1987) "Location theory, regional science, and economics," *Journal of Regional Science*, Vol. 27, pp. 519-528.

Toivanen, O. and Niininen, P. (1998) "Investment, R&D, subsidies and credit constraints," Working Paper, Department of Economics MIT and Helsinki School of Economics.

Trochim, W. (2000) *The research methods knowledge base*, 2nd Edition, accessible at Internet WWW page, at URL: <http://trochim.human.cornell.edu/kb/index.htm>

Utterback, J. (1994) *Mastering the dynamics of innovation*, Cambridge, MA, Harvard Business School Press.

Vernon, R. (1966) "International investment and international trade in the product cycle," *Quarterly Journal of Economics*, Vol. 80, pp. 190-207.

Vernon, R. (1979) "The product cycle hypothesis in a new international environment," *Oxford Bulletin Journal of Economics*, Vol. 41, pp. 255-267.

Vigdor, J. (1998) "Was Tiebout wrong? Evidence from proposition 2 1/2 in Massachusetts," Harvard University mimeograph.

Von Hipple, E. (1994) "Sticky information and the locus of problem solving: Implications for innovation," *Management Science*, Vol. 40, pp. 429-439.

Von Tunzelmann, N. and Martin, B. (1998) "Public vs. private funding of R&D and rates of growth: 1963–1995," Working Paper, Science Policy Research Unit, University of Sussex.

Wagner, J. (2001) "Regional economic diversity: Action, concept, or state of confusion," *Journal of Regional Analysis and Policy*, Vol. 30, pp. 1-22.

Wallsten, S. (2001) "The role of government in regional technology development: the effects of public venture capital and science parks," *Stanford Institute for Economic Policy Research*, Discussion Paper No. 00-39.

Wasylenko, M (1997) "Taxation and economic development: The state of economic literature," *New England Economic Review*, pp. 37-52.

Weber, A. (1909) *Theory of the location of industries*: Translated by Friedrich, C. (1929), Chicago: University of Chicago Press.

White, H. (1980), “A heteroskadasticity-consistent covariace matrix estimator and a direct test for hetroskedasticity,” *Econometrica*, Vol.48, pp. 817-838.

Wooldridge, J. (2003) *Introductory econometrics: A modern approach,* 2nd edition, United States, Thomson- southwestern.

Wundt, B. (1992) ‘‘Reevaluating alternative measures of industrial diversity as indicators of regional cyclical variations,’’ *Review of Regional Studies*, Vol. 22, pp. 59-73.

Yelowitz, A. (1995) “The Medicaid Notch, labor supply and welfare participations: Evidence from eligibility expansions,” *The Quarterly Journal of Economics*, Vol. 110, pp. 909-939.

Zipf, G. (1949) *Human behavior and the principle of least effort*, Cambridge, MA, Addison–Wesley Press.