ESSAYS ON LABOR AND PRODUCT MARKETS

by

Hakki Lee

B.A. in Economics, Kyunghee University, 2007

M.A. in Economics, Korea University, 2009

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences in partial

fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Hakki Lee

It was defended on

April 29, 2016

and approved by

Dr. Daniele Coen-Pirani, University of Pittsburgh, Economics

Dr. Marla Ripoll, University of Pittsburgh, Economics

Dr. James Cassing, University of Pittsburgh, Economics

Dr. Todd Schoellman, Arizona State University, W.P. Carey School of Business

Dissertation Advisors: Dr. Daniele Coen-Pirani, University of Pittsburgh, Economics,

Dr. Marla Ripoll, University of Pittsburgh, Economics

ESSAYS ON LABOR AND PRODUCT MARKETS

Hakki Lee, PhD

University of Pittsburgh, 2016

This dissertation is a collection of three essays on labor economics and industrial organization. In the first essay, I study how and to what extent labor market frictions, which are defined as the inability to costlessly find jobs or move between jobs, affect workers' schooling decisions. In order to study this link, I use an on-the-job search model. Consistent with the data, the model economy predicts that a higher job-to-job transition rate increases the opportunity cost of a college education, reducing the incentives for schooling. Instead, a higher job separation rate decreases the opportunity cost, leading to more schooling. In addition, a higher job finding rate increases employment duration, which can help college educated workers enjoy higher skill prices for a longer time and lead to more investment in schooling.

In the second essay, I explore whether the number and types of people who attend college affect the college wage premium, which is defined as the wage of college graduates relative to high school graduates. I estimate the return to schooling by exploiting all of the variation or only the exogenous variation in educational attainment as the instrumental variable. The result reveals that composition effects account for nearly one half of the observed return to schooling. In addition, I find strong evidence that sudden short-run changes in educational attainment lowered education wage premiums for the affected cohorts, but only weak evidence that long-run trend changes in educational attainment lowered education wage premiums.

The third essay analyzes a monopolistic supplier's optimal decision of input prices when two downstream sellers simultaneously choose their advertisement efforts and their output levels. The independent advertisement decision by each seller causes the free-rider problem by its rivals. The supplier uses price-discrimination to alleviate the free-riding problem associated with the advertisement decision. Therefore, allowing price-discrimination may increase aggregate output– and social surplus– the reverse of the welfare implications in the previous literature.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This work was shaped by the advice and support of many people over several years. I am very grateful to my dissertation committee members: Daniele Coen-Pirani, Marla Ripoll, James Cassing, and Todd Schoellman. I very much appreciate their guidance, motivation and insightful feedback on my research. I am also grateful for the many hours that Sewon Hur has encouraged and supported me as an excellent mentor.

Finally, I want to thank my parents, Byeongkyu Lee and Suin Park, and my sister, Jungeun Lee, for their great support and care. I am also deeply indebted to Zeynep Kabucouglu for her positivity and encouragement.

1.0 INTRODUCTION

This dissertation is a collection of three essays on labor economics and industrial organization. The first two essays focus on individuals' schooling choices and their effect on the college wage premium. In the final essay, I direct my attention to another important issue in the product market. I focus on the merits of price discrimination in input markets, a topic that has long been debated intensively. In particular, I study the effects of third-degree price discrimination in input markets when downstream firms can enhance demand by advertising.

The first essay, titled *"Schooling Investment in Frictional Labor Markets,"*[1] is presented in Chapter 2. In this paper, I address the question of how and to what extent labor market frictions, which are defined as the inability to costlessly find jobs or move between jobs, affect individuals' schooling decisions. First, I present some empirical evidence that there is a correlation between labor market frictions and college attainment across countries. The literature has focused on various aspects of human capital investment, but not on the relationship between labor market frictions and college attainment, which has yet to receive much attention. In order to investigate the channels of the relationship between labor market frictions and individuals' schooling decisions, I develop an on-the-job search model with workers' schooling choice prior to labor market entry, in which firms are heterogeneous in their productivities and workers are heterogeneous with respect to the amenity value of schooling. Before labor market entry, workers decide whether to go to college by taking into account schooling costs and the difference between the value of unemployment for high school graduates and college graduates. High school graduates experience multiple job turnovers toward better paying jobs, while their counterparts spend years in college. If they move up the wage ladder through job turnover during their first four years in the labor market, the

[1] This research is a joint work with Seung-Gyu Sim.

opportunity cost of college education during this period may not be negligible. Therefore, considering the opportunity cost of a college education is an important channel.

The main finding of this paper is that higher job finding rates or separation rates can increase schooling, while an increase in the job-to-job transition rate discourages schooling. This finding is consistent with the empirical evidence. The key intuition for this result is that when the job finding rate decreases, the unemployment duration increases, which means that the period in which a worker exploits his human capital shortens. Therefore, the shorter duration leads to a lower college attainment rate. If the job separation rate increases, then the unemployment duration gets longer and the opportunity cost of a college education decreases, both of which have opposite effects on schooling. First, a worker faces less time in which he can exploit his human capital on the job, so his college attainment may decrease. Second, a worker with a high school degree faces a higher chance of being separated from his job, which lowers the opportunity cost of choosing a college education. This leads workers to get more education before entering the labor market. As a result, when the latter effect dominates the former one, college attainment increases. Finally, a lower job-to-job transition rate implies that a worker with only a high school degree cannot easily move to jobs that pay better while his counterpart is in college. Hence, the opportunity cost of getting a college education is lower, which induces an increase in the college attainment rate.

Under the presence of labor market frictions, I address the interesting question of whether the current level of education investment is socially efficient. To answer this question, I assume that there is a benevolent social planner who cares only about the welfare of workers and who decides how workers will choose schooling and how they are matched with jobs, taking labor market frictions as given. Surprisingly, I find that there is an over-investment in schooling. The intuition behind this result is that firms tend to suppress unskilled wage offers, and the wage differentials exceed productivity differentials. As a result, returns to higher education get bigger in the market equilibrium.

In Chapter 3, titled *"Policy-Induced Variation in College Labor Supply and the Skill Premium,"*[2] I test for the importance of composition effects in affecting levels and changes of education wage premiums in the context of Korea. Korea offers an interesting case because

[2]This research is a joint work with Jaehan Cho and Todd Shoellman

it has experienced a rise in educational attainment and education wage premiums, like most other advanced economies. However, the rise in education in Korea was much larger and more rapid than that in most other advanced economies. Further, some of the increase was policy-induced.

In this paper, I adopt two empirical strategies previously used in the literature to quantify ability and composition effects in the United States in a consistent and unified way. My analysis yields two main results. The first concerns the importance of composition effects in accounting for the cross-sectional return to schooling. In line with Kaymak (2009), I find that composition effects account for nearly one half of the observed return to schooling, indicating that the true private return to schooling is slightly more than half of the observed Mincer return. The second result concerns the importance of composition effects in the time series. Consistent with Juhn et al. (2005), I find evidence that increasing college attainment lowers the college wage premium for a cohort, suggesting that higher college attainment is obtained by lowering the relative ability of college graduates.

Chapter 4 presents the third essay, titled *"Input Price Discrimination, Informative Advertisement Efforts, and Welfare."*[3] The paper investigates optimal linear wholesale pricing by a monopolistic manufacturer that serves two downstream firms. The downstream firms differ in production costs, produce a homogeneous product, and compete in quantities. The manufacturer may be constrained by a non-discrimination clause, and I analyze the effects of such a legal rule. The novelty of the paper is in allowing downstream firms to exert costly effort in order to boost market demand (interpreted as advertising expenditures). Downstream firms' effort choices are substitutes, and, thus, a free-riding problem arises. Moreover, because effort choices are non-contractible, the wholesale prices are the only available instrument to incentivize downstream firms.

A well-known result in the literature on price discrimination in intermediate goods markets, as in Katz (1987), DeGraba (1990), and Yoshida (2000), is that the less efficient firm receives a discount. This immediately implies that permitting price discrimination reduces allocative efficiency–i.e., output is shifted to the less efficient firm–and, for linear demand, also reduces total welfare. If downstream firms can invest in demand-enhancing efforts, these

[3]This research is a joint work with Seung-Gyu Sim.

findings may be reversed, as this paper shows: For certain parameter ranges, the more efficient firm receives a discount, and permitting price discrimination increases total welfare. The main insight is that under price discrimination, the supplier can alleviate the free-riding problem with regard to advertising efforts by offering the more efficient downstream firms a discount, which also furthers allocative efficiency. In particular, the supplier has an incentive to discriminate between downstream firms even if these are identical. Numerical simulations show that this reduction in the free-riding problem can be strong enough to make a ban on price discrimination harmful to welfare.

2.0 SCHOOLING INVESTMENT IN FRICTIONAL LABOR MARKETS[1]

2.1 INTRODUCTION

It is not surprising that labor market frictions, which are defined as the inability to costlessly find jobs or move between jobs, affect workers' incentives to acquire skills and accumulate human capital. Indeed, numerous papers study human capital accumulation through learning-by-doing or continuous investment by either employers or employees in a frictional labor market.[2] However, there are not many papers regarding how labor market frictions affect workers' schooling decisions prior to their labor market entry. It is well known that schooling is the most efficient way of accumulating human capital. A simple illustration that shows there is a relationship between college attainment rates and the labor market frictions is presented in Figure 1, using partial regressions.[3] It depicts that college attainment rate is correlated positively with the job finding and separation rates (but weaker correlation), and negatively correlated with job turnover rate. The regression analysis also reports that job finding, separation, and turnover rates have (95 percent confidence level) significant partial effects by themselves. Motivated by Figure 1, this paper separately identifies the channels through which labor market friction affects returns to higher education and its opportunity cost, particularly focusing on the joint influence of these channels. In addition, it analyzes

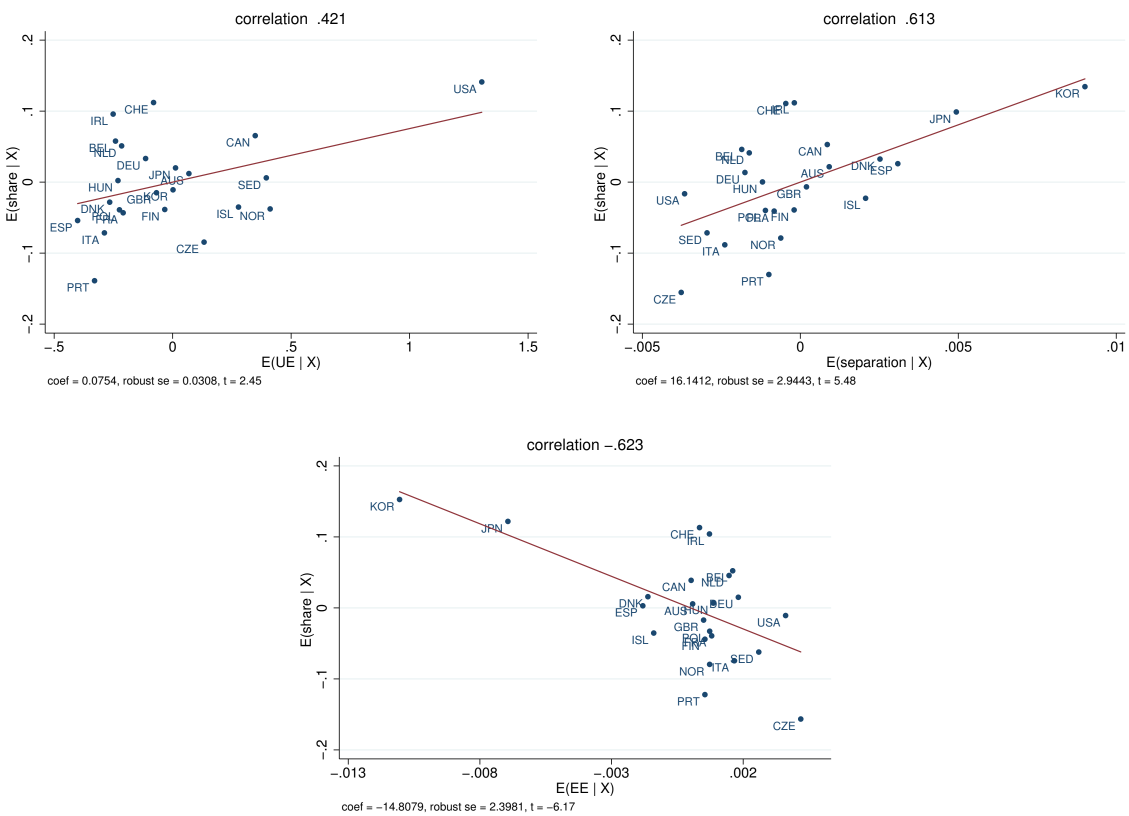
[1]This research is a joint work with Seung-Gyu Sim.

[2]See Acemoglu (1997) and Acemoglu and Pischke (1998) for on-the-job training and Burdett et al. (2011) and Bagger et al. (2014) for learning-by-doing on the job.

[3]We obtain the partial regression estimates using the following procedure. For each plot, first, we regress college attainment rate on the isolated labor market friction that is of interest. Next, the fitted college attainment rate is regressed on the fitted labor market friction. This analysis is performed using a cross-sectional data that consist of 27 OECD countries in total, which are Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden Switzerland, the United Kingdom, and the United States.

the efficiency of the decentralized market equilibrium.

Figure 1: Search Friction and College Attainment Rate



Note: We present three panels that show partial correlations between college attainment rates (share) and job finding rate (UE), employment-to-employment transition rate (EE) and separation rate (separation), respectively. The control variables (X) change by the labor market friction rate of interest. For instance, in the upper left panel, the y-axis shows the expected college attainment rate conditional on employment-to-employment transition rate, separation rate and education spending. The x-axis shows the expected job finding rate conditional on employment-to-employment rate, separation rate and education spending. The college attainment rates are borrowed from Barro and Lee (2013) and its companion webpage `http://www.barrolee.com/`. The sample size is limited to the available data for the labor market frictions. The information for the worker flows are mostly borrowed from Hobijn and Şahin (2009) and it is available for 27 countries. Korean data are calculated from the Korean Labor and Income Panel Study.

In reality, workers enter the labor market as unemployed. Once they become employed, they start working and share the match surplus with their employers according to a certain wage norm. Then, they repeatedly experience unemployment and employment until they retire. Given mortality of human beings, Manuelli and Seshadri (2009), Córdoba and Ripoll (2013), and Restuccia and Vandenbroucke (2014) argue that expected lifespan is one of the most important factors when accounting cross country schooling differences. Apart from life expectancy, an individual worker's decisions pertaining to education and schooling are related to employment spells, skill prices, and the presence of search frictions. Consequently, higher skill prices for longer utilization spells encourage investment in schooling. Another significant channel through which labor market friction affects workers' schooling decision prior to their labor market entry is the opportunity cost of schooling. Topel and Wald (1992) find that a typical worker holds seven jobs during the first 10 years of employment. Sim (2014) reports the average job duration of American white male high school graduates to be slightly over two years. Those stylized facts imply that high school graduates experience multiple job turnovers (toward better paying jobs), while their counterparts in college have opportunity costs during school that may not be negligible.

To capture those channels, this paper incorporates workers' schooling decisions into the Burdett and Mortensen (1998) framework, which is a canonical framework in the literature of on-the-job search.[4] Firms are heterogenous in their productivities. Workers are heterogenous with respect to the amenity value of (or disutility from) schooling.[5] Workers enter the labor market either after high school or college graduation; firms post skill-dependent wages. Firms know the labor force available for the job evolves with the employee's wage. Firms that can offer higher wages can keep the worker for a longer time at the steady-state. Also, they can hold a larger work force because they will attract more workers to the company. Given the heterogeneity in the education of workers, firms compete with each other to attract high-skilled workers because college educated workers can contribute to the productivity of the firm more than high school graduates. However, wage offer distributions, rarely observed in general, are expected to be first order stochastically dominated by the wage earning distributions due to on-the-job search in frictional labor markets.[6] By developing a

[4]Shimer (2006) points out that the axiomatic Nash bargaining solution, adopted in usual wage-bargaining models, is inapplicable to a search and matching model with on-the-job search because the set of feasible payoffs is typically nonconvex so that an increase in the wage raises the duration of an employment relationship.

[5]The assumption of heterogenous schooling values is common in labor and macro literature, for example Heckman et al. (1998a), Bils and Klenow (2000), and Restuccia and Vandenbroucke (2013a).

[6]Christensen et al. (2005), using data from the Danish Integrated Database for Labour Market Research, estimate a job search model with on-the-job search and separately identifies the wage offer distribution from the wage earning distribution.

structural model, we elicit the opportunity cost based on the wage offer distribution.

The main finding of this paper is that labor market frictions have an important role in schooling decisions through two channels: employment duration and the opportunity cost of schooling. It shows that the employment duration for college graduates rises with the job finding rate and declines with the job separation rate as well as the retirement rate. Therefore, higher employment duration allows college graduates to enjoy higher skill prices for a longer time. The opportunity cost of higher education increases in the employment-to-employment transition rate and decreases in the job separation rate. Consistent with the empirical findings, the calibrated model based on the U.S. labor market data predicts that an increase in job finding or separation rate encourages schooling but an increase in employment-to-employment transition rate discourages schooling. The main intuition behind these results is the following: When the job finding rate decreases, the unemployment duration increases, which means that the period a worker exploits the human capital shortens. Therefore, the shorter duration leads to a lower college attainment rate. If job separation rate increases, then longer unemployment duration and lower opportunity cost of choosing college education is obtained, both of which have opposite effects on schooling. First, a worker faces lesser time that he can exploit his human capital on the job, so the college attainment may decrease. Second, a worker with a high school degree faces higher chance to be separated from the job. Therefore, there is a lower probability to move up on the wage ladder. This leads workers to want to pursue more education before entering the labor market. As a result, the latter effect dominates the former one, so a higher college attainment rate is obtained. Last, when the employment-to-employment transition rate is lower, it implies that a worker with only a high school degree cannot easily move up on the wage ladder while his counterpart is in college. Hence, the opportunity cost of getting a college education is lower, which induces an increase in college attainment rate.

For efficiency analysis, we develop the constrained social planner problem, who takes labor market friction as given and makes schooling decisions on behalf of individual workers in order to maximize the present value of the expected net output flow throughout their career, as in Sim (2015). Our numerical experiments comparing the market equilibrium outcome with the planner's outcome show that firms suppress unskilled wage offers further than skilled wage offers overall. This makes the (discounted) skilled-unskilled wage differentials exceed (discounted) productivity differentials and escalates returns to higher education in the market equilibrium. Recently, Flinn and Mullins (2015) extend the wage-bargaining model to allow for endogenous schooling choice. In their model, the weak bargaining position of workers, which is assumed to be same across different skill levels, reduces wage differentials transmitted from productivity differentials. As a result, there is an under-investment in schooling. In these models, there is also the standard holdup problem that emerges because wage contracts are not determined at the time the schooling decision is made. Had we assumed wage determination by Nash bargaining, the holdup problem might be exacerbated particularly when the workers' bargaining power is low. This would lead to a larger mismatch between aggregate returns to schooling and workers' private returns. However, our model abstracts from a match function and Nash bargaining, so holdup is relatively not an issue. In contrast, we, by endogenizing the wage offer distributions, demonstrate that the unskilled wage offers, being further suppressed than the skilled offers, result in over-investment in schooling.[7]

Flinn and Mullins (2015), concerning under-investment in schooling, propose that schooling subsidies improve aggregate welfare. Although our paper indicates over-investment in schooling, it suggests that enhancing the quality of higher education improves social welfare. It shows that the over-investment problem (relative to the constrained planner's outcome) can be resolved by encouraging college attainment because the college enrollment decision by individual workers in the decentralized market equilibrium is less elastic to the changes in education cost and quality than the enrollment decision by the constrained social planner in the centralized efficiency benchmark. For example, if the college graduates' average productivity improves, the planner accounts for the full improvement in determining the college enrollment rate, whereas those individual workers in the market equilibrium consider the improvement in their expected wage offers, which partially reflects the productivity improvement. This counterfactual experiment suggests that the U.S. government can resolve the over-investment problem by encouraging the college attainment rate through lowering

[7] Charlot and Decreuse (2005) also point out that over-education may take place in a frictional labor market because of matching externality and self-selection.

education cost and improving education quality.

The rest of the paper proceeds as follows: Section 2.2 provides empirical results that shows the relationship between college attainment and labor market frictions; Section 2.3 develops the theoretical model and characterizes the steady state equilibrium; Section 2.4 presents the numerical analysis; Section 2.5 further discusses implications of our study; and Section 2.6 concludes.

2.2 EMPIRICAL MOTIVATION

In this section, we provide empirical results that highlight the relationship between schooling and labor market frictions. We use the following specification in this empirical analysis:

$$Schooling_j = \alpha_0 + \alpha_1 f_j + \alpha_2 EE_j + \alpha_3 s_j + \alpha_4 spending_j + \varepsilon_j, \tag{2.1}$$

where f_j is job finding rate, EE_j is employment-to-employment transition rate, s_j is separation rate and $spending_j$ is education spending for each country j. We use college attainment rate as a measure of schooling. The college attainment rates are borrowed from Barro and Lee (2013). The data for worker flows, such as job finding rate, employment-to-employment transition rate, and separation rate, are used as measures of labor market frictions. Most of the information about worker flows for all countries except Korea is from Hobijn and Şahin (2009). Korean data are calculated from the Korean Labor and Income Panel Study (KLIPS).[8] We also control for education spending.[9]

Table 1 shows the estimation results of equation (2.1). The coefficients for job finding rate and separation rate are positive and significant in all specifications. This relationship can be explained by the following: As job finding rate increases, it shortens unemployment spells. This gives more incentives to workers to invest in college education so that they can benefit from higher returns to schooling for a longer period of time. The positive correlation between college attainment rate and separation rate can be explained by the opportunity cost. As

[8] See Appendix A.1 for more-detailed information about the sample and the procedure of computation.

[9] The data source is OECD (2015), Education spending (indicator). For more detailed information, see https://data.oecd.org/eduresource/education-spending.htm.

Table 1: Cross-country regressions explaining schooling with search frictions

	Dependent Variable: College Attainment Rate.			
Independent Variable	(1)	(2)	(3)	(4)
Job finding rate	0.072**	0.086**	0.063*	0.075**
	(0.032)	(0.033)	(0.033)	(0.033)
Job-to-job transition rate	–	−1.434	−13.900***	−14.808***
		(1.260)	(1.843)	(2.590)
Separation rate	–	–	14.310***	16.141***
			(2.683)	(3.180)
Education spending	–	–	–	0.001
				(0.180)
No. of countries	27	25	25	23
R^2	0.10	0.17	0.42	0.52

note: The college attainment rates are borrowed from Barro and Lee (2013) and its companion webpage http://www.barrolee.com/. The information about the worker flows is mostly borrowed from Hobijn and Şahin (2009). Korean data are calculated from the Korean Labor and Income Panel Study (KLIPS). Estimation is by robust standard errors. Standard errors are reported in parentheses. (∗∗∗, ∗∗, ∗ represent significance level at 1%, 5% and 10%, respectively.)

separation rate increases, the opportunity cost of getting a college education becomes smaller as the value of forgone earnings had workers chosen going to labor market instead of going to college decreases. Finally, the results show that employment-to-employment transition rate is negatively correlated with college attainment rate. The intuition behind this result is related to the increasing opportunity cost. As high school graduates move up on the wage ladder faster, the value of forgone earnings increases, which hampers the incentives to invest in college education. These results imply that the labor market frictions play an important role in explaining the educational attainment across countries.

One caveat is that it is difficult to measure the true effect of each labor market friction on college attainment using this simple regression analysis. However, we can compare each coefficient in order to get an intuition about the relative importance of each factor. In the sample, the standard deviation of job finding rate is 0.36. The coefficient of job finding rate on college attainment rate is estimated as 0.072. As a result, if the job finding rate increased by one standard deviation, the college attainment rate would increase by about 2.5

percent. Similarly, since the standard deviation of separation rate is 0.01 and the coefficient is estimated as 16.14, the college attainment rate would increase by 19.2 percent when there is one standard deviation increase in the separation rate. As for the employment-to-employment transition rate, the sample standard deviation is 0.01. The coefficient of the employment-to-employment transition rate is estimated as -14.81, which implies that if there were one standard deviation increase in the employment-to-employment transition rate, then the college attainment rate would decrease by 19.6 percent.

2.3 MODEL

2.3.1 Environment

We extend the job search model with on-the-job search proposed by Burdett and Mortensen (1998) by introducing schooling choices before entering the labor market. Time is continuous with an infinite horizon and only steady states are considered for tractability. There is a continuum of firms and workers, each of measure one. Firms are immortal and heterogenous in terms of their productivity. Let us denote by $H(p)$ the proportion of firms whose productivity is less than p. It is assumed that $H(p)$ is continuously differentiable and has a finite support of $[\underline{p}, \overline{p}]$. Workers are mortal and leave the labor market at rate ρ. Each retiree is replaced with a newly born worker, who makes a schooling decision before entering the labor market. The worker without a college degree, called 'unskilled worker' ($i = 0$) hereafter, produces $p + s_0$ when she is employed at the firm with productivity p, but the worker with a college degree, called 'skilled worker' ($i = 1$), produces $p + s_1$ at the same firm. Both firms and worker discount future at rate r.

The newly born worker decides whether to enter the labor market as an unskilled worker or as a college-educated worker before entering the labor market. Once she decides to go to college, she will enter the labor market as a skilled worker after college. The idiosyncratic preferences for higher education, ε, is assumed to follow a (standard) logistic distribution, $Logistic(\mu, \sigma)$. It is assumed that $\mu < 0$ due to college tuition, living cost and so on. Then, the implied college enrollment probability is given by

$$\psi = [1 + \exp[-(\exp(-\Delta(r+\rho))U_1 - U_0 + \mu)/\sigma]]^{-1}, \tag{2.2}$$

where Δ represents the duration for college education and (U_0, U_1) are the expected values of unemployment without and with a college degree, respectively. We borrow the schooling decision function from Ishimaru et al. (2015).

All individuals are either unemployed or employed. If a worker is unemployed, she collects unemployment benefit b and has a chance to get a job with an exogenous contact rate λ_U. The lifetime value of an unemployed worker, whether she has a college degree or not, is given by

$$rU_i = b - \rho U_i + \lambda_U \int \max\{W_i(x) - U_i, 0\} dF_i(x) \quad \text{for each} \quad i \in \{0, 1\}, \tag{2.3}$$

where F_i is the cumulative distribution of offered wages and $W_i(x)$ represents the lifetime value of the employed worker receiving wage x. The left-hand side of equation (2.3) represents the opportunity cost of holding asset 'i-type unemployment,' and the right-hand side indicates the flow value of unemployment benefit, the losses from retirement, and the net gains from job-finding.

The employed worker earning ω_i faces layoff risk at rate δ, and gets an offer from another job with an exogenous contact rate λ_E,

$$rW_i(\omega_i) = \omega_i - \rho W_i + \delta(U_i - W_i(\omega_i)) + \lambda_E \int \max\{W_i(x) - W_i(\omega_i), 0\} dF_i(x). \tag{2.4}$$

A job operated by an unskilled worker accrues revenues $p + s_0$ to the firm and pays ω_0 to the worker, and a job with a skilled worker accrues $p + s_1$ and pays w_1 to the workers, where p is firm's productivity and s_i is the worker's productivity ($s_1 > s_0$). Let $J_i(\omega_i, p)$ be the firm's value of the match producing $p_i + s_i$ and paying ω_i, given by

$$rJ_i(\omega_i, p) = p + s_i - \omega_i - [\rho + \delta + \lambda_E(1 - F_i(\omega_i))] J_i(\omega_i, p), \quad \text{for each} \quad i \in \{0, 1\}. \tag{2.5}$$

Each firm posts a vacancy at every instant. Once the vacancy is matched with a worker, it immediately starts producing and pays the pre-committed skill-dependent wage. Let $\{u_0, u_1\}$ denote the masses of unskilled and skilled unemployed workers, respectively. In addition, let $\{G_0(\omega_0), G_1(\omega_1)\}$ be the proportion of skilled and unskilled workers with wages no greater than $\{\omega_0, \omega_1\}$, respectively. The firm with productivity $p \in [\underline{p}, \overline{p}]$ optimally posts $\{\omega_0(p), \omega_1(p)\}$ to maximize the following expected value

$$\sum_{s=0,1} \Big[(\lambda_U u_i + \lambda_E G_i(\omega_i(p))) J_i(\omega_i(p), p) \Big]. \tag{2.6}$$

The first-order-conditions with respect to $\omega_0(p)$ and $\omega_1(p)$ implies

$$\frac{\lambda_U u_i + \lambda_E G_i(\omega_i(p))}{\lambda_E J_i(\omega_i(p), p)} = \frac{\lambda_U u_i + \lambda_E G_i(\omega_i(p))}{r + \rho + \delta + \lambda_E(1 - F_i(\omega_i(p)))} \frac{\partial F_i(\omega_i(p))}{\partial \omega_i} + \frac{\partial G_i(\omega_i(p))}{\partial \omega_i}, \tag{2.7}$$

for each $i \in \{0, 1\}$. Given workers' on-the-job search behaviors, it is natural to think that a more productive firm offers a higher value so that it will not allow any less productive poaching firms on equilibrium. It implies that

$$F_i(\omega_i(p)) = H(p) \text{ and } \frac{\partial H}{\partial p} = \frac{\partial F_i(\omega_i)}{\partial \omega_i} \frac{\partial \omega_i(p)}{\partial p} \text{ for each } i \in \{0, 1\}. \tag{2.8}$$

Given firms' productivity distribution $H(p)$, the steady state equilibrium with schooling decision and on-the-job search consists of the value function $\{W_i, U_i, J_i\}_{i=0,1}$ and steady state measures $(\{F_i, u_i, G_i\}_{i=0,1})$, which jointly satisfies the following conditions:

(i) Given $\{F_i\}_{i \in \{0,1\}}$, workers make an optimal decision regarding job turnover from equations (2.3) and (2.4), which jointly determine $\{U_i, W_i\}_{i=\in\{0,1\}}$.

(ii) Given $\{U_i\}_{i \in \{0,1\}}$, the newly born workers make schooling decisions prior to the labor market entry through (2.2).

(iii) Given $H(p)$ and $\{F_i\}_{i=\in\{0,1\}}$, each recruiting firm with productivity p optimally posts a pair of wages $(\omega_0(p), \omega_1(p))$ such that it satisfies the first order condition (2.7), which determines $\{J_i\}_{i \in \{0,1\}}$ through (2.5).

(iv) $\{F_i, u_i, G_i\}_{i \in \{0,1\}}$ are uniquely determined.

2.3.2 Characterization of the Steady State Equilibrium

Let us first consider the steady state worker flow and implied distributions. In the steady state equilibrium,

$$(\rho+\lambda_U)u_0 = \delta G_0(\omega_0(\overline{p})) + \rho(1-\psi) \;\text{ and }\; (\rho+\delta)G_0(\omega_0(\overline{p})) = u_0\lambda_U. \tag{2.9}$$

It implies that

$$u_0 = \frac{(\rho+\delta)(1-\psi)}{(\rho+\delta+\lambda_U)} \;\text{ and }\; G_0(\omega_0(\overline{p})) = \frac{\lambda_U}{\rho+\delta}u_0. \tag{2.10}$$

Let $s(t)$ be the steady state measure of college students who have spent $t < (\Delta)$ years in schooling. Equating the inflow into and the outflow from the pool of 'students,' yields that

$$\rho\psi dt = (1-\rho dt)(s(\Delta)-s(\Delta-dt)) + \rho dt s(\Delta-dt) + o(dt^2). \tag{2.11}$$

The left-hand side represents the inflow and the right-hand side represents the outflow due to graduation toward labor force and early retirement toward non-labor force. Dividing both sides by dt, sending dt to zero, and solving the differential equation yields

$$s(\Delta) = \psi - \psi\exp(-\rho\Delta). \tag{2.12}$$

Since $u_1 + s(\Delta) + G_1(\omega_1(\overline{p})) = \psi$ and $\lambda_U G_1(\omega_1(\overline{p})) = (\rho+\delta)u_1$, we obtain

$$u_1 = \frac{(\rho+\delta)\psi\exp(-\rho\Delta)}{\rho+\delta+\lambda_U} \;\text{ and }\; G_1(\omega_1(\overline{p})) = \frac{\lambda_U}{\rho+\delta}u_1. \tag{2.13}$$

The unemployed worker finds a job with wage payment no greater than ω_i at rate $\lambda_U F_i(\omega_i)$. The worker working at a job with wages less than ω_i switches to a higher valued job at rate $\lambda_E(1-F_i(\omega_i))$, is laid off at rate δ, and retires at rate ρ. The steady state measure $G_i(\omega_i)$ is characterized by

$$\lambda_U F_i(\omega_i)u_i = (\rho+\delta+\lambda_E(1-F_i(\omega_i)))G_i(\omega_i). \tag{2.14}$$

Taking derivative of (2.14) with respect to ω_i and reordering yields

$$\frac{\partial G_i(\omega_i)}{\partial\omega_i} = \frac{\lambda_u u_i + \lambda_E G_i(\omega_i)}{\rho+\delta+\lambda_E(1-F_i(\omega_i))}\frac{\partial F_i(\omega_i)}{\partial\omega_i}, \quad \text{for each } \; i\in\{1,2\}. \tag{2.15}$$

Finally, solving differential equation (2.15) yields

$$G_i(\omega_i) = \frac{1}{I_i(\omega_i)} \int_{\underline{\omega}_i}^{\omega_i} \frac{I_i(x)\lambda_U f_i(x) u_i}{\rho + \delta + \lambda_E(1 - F_i(x))} dx, \tag{2.16}$$

where

$$I_i(x) = \exp\Big[\int_{\underline{\omega}_i}^{x} \frac{-\lambda_E f_i(z)}{\rho + \delta + \lambda_E(1 - F_i(z))} dz\Big]. \tag{2.17}$$

In equilibrium, the least productive firm should attract only unemployed workers. It optimally posts $(W_0(\omega_0(\underline{p})), W_1(\omega_1(\underline{p}))) = (U_0, U_1)$. Lemma 1, together with $F_0(\omega_0(\underline{p})) = F_1(\omega_1(\underline{p})) = G_0(\omega_0(\underline{p})) = G_1(\omega_1(\underline{p})) = 0$, determines the initial values of the system of differential equations.

Lemma 1 *Suppose that (U_0, U_1) are given. Equation (2.2) determines ψ. The least productive firm posts*

$$\omega_i(\underline{p}) = (r + \rho)U_i - \frac{\lambda_E}{\lambda_U}[(r + \rho)U_i - b], \quad \textit{for each} \quad i \in \{0, 1\}, \tag{2.18}$$

which results in

$$\begin{aligned} W_i(\omega_i(\underline{p})) &= \frac{\omega_i(\underline{p}) + \delta U_i}{r + \rho + \delta} + \frac{\lambda_E((r + \rho)U_i - b)}{(r + \rho + \delta)\lambda_U}, \quad \text{and} & (2.19) \\ J_i(\omega_i(\underline{p}), \underline{p}) &= \frac{\underline{p} + s_i - \omega_i(\underline{p})}{r + \rho + \delta + \lambda_E} \quad \text{for each} \quad i \in \{0, 1\}. & (2.20) \end{aligned}$$

Given the initial values (2.18)-(2.20) and $G_i(\omega_i(\underline{p})) = 0$, the steady state equilibrium is characterized by the initial value problem (IVP) having the following system of differential equations.

$$\begin{aligned} \frac{dW_i(\omega_i)}{d\omega_i}\frac{d\omega_i(p)}{dp} &= \frac{1}{r + \rho + \delta + \lambda_E(1 - H(p))}\frac{d\omega_i(p)}{dp}, & (2.21) \\ \frac{dJ_i(\omega_i(p), p)}{dp} &= \frac{1}{r + \rho + \delta + \lambda_E(1 - H(p))}\Big[1 - \frac{d\omega_i(p)}{dp} + \lambda_E J_i(\omega_i, p)\frac{dH(p)}{dp}\Big], & (2.22) \\ \frac{dG_i(\omega_i(p))}{d\omega_i}\frac{d\omega_i(p)}{dp} &= \frac{\lambda_U u_i + \lambda_E G_i(\omega_i)}{\rho + \delta + \lambda_E(1 - H(p))}\frac{dH(p)}{dp}, \quad \textit{and} & (2.23) \\ \frac{d\omega_i(p)}{dp} &= \lambda_E J_i(\omega_i(p), p)\Big[1 + \frac{r + \rho + \delta + \lambda_E(1 - H(p))}{\rho + \delta + \lambda_E(1 - H(p))}\Big]\frac{dH(p)}{dp}. & (2.24) \end{aligned}$$

Lemma 1 says that the whole system can be pinned down at a fixed point problem in the $\mathbb{R} \times \mathbb{R}$ space. Taking (U_0, U_1) as given, we get the whole description of the steady state equilibrium. In equilibrium, the lifetime values of (U_0, U_1) should be consistent with (2.3). Since the analytical proof of the existence and uniqueness of the result is somewhat challenging, we replace it with numerical experiments with different initial guesses in Section 3. This section ends with the following Lemma 2.

Lemma 2 *Given* $p \in [\underline{p}, \overline{p}]$, $\frac{d\omega_0(p)}{dp} < \frac{d\omega_1(p)}{dp}$ *if and only if* $J_0(\omega_0(p), p) < J_1(\omega_1(p), p)$.

To understand the implications of Lemma 2, it is important to recognize the key properties of the model. In the model presented, once the agents are in the labor market, employed workers randomly search for jobs that offer higher wages than their existing jobs. Unemployed workers look for jobs that offer more than their reservation wages. The firms, taking into account the search characteristics and the wage distribution in the market, post their own wages. Burdett and Mortensen (1998) show that in a simpler job-search model, there is an equilibrium in which the economy supports a variety of wages for workers that are equally productive. Our model is similar in the sense that, for each education group, the firm knows that the labor force available for the job evolves with the employers wage. Firms that can offer higher wages can keep the worker for a longer time at the steady-state. Also, they can hold a larger work force because they will attract more workers to the company. Given the heterogeneity in the education of workers, firms compete with each other to attract the high-skilled workers because the college educated workers can contribute to the productivity of the firm more than the high school graduates. Lemma 2 shows that in order to keep the skilled workers, the wages posted will be higher particularly by those firms with higher productivity. By doing so, firms award the skilled worker to attract them. This implies that the increase in wages that unskilled worker faces is quite lower during his career path and generates a more suppressed wage relative to the wages of the skilled worker in the market equilibrium. In other words, highly productive firms want to differentiate themselves and prevent low productive firms from stealing their workers by offering high wages. This will generate a large pool of skilled workers working for high productivity firms in this economy at the steady-state.

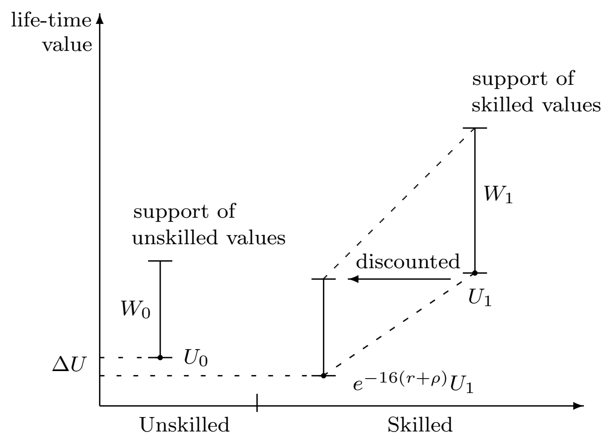
Let us also explain the mechanisms that are at play regarding the workers' problem. The aforementioned firms' equilibrium strategy can help a college educated worker move up the wage ladder faster compared to a high school graduate given the labor market conditions. This generates higher incentives for college education among agents making an education-related decision. However, there is also the standard holdup problem that emerges in an economy with labor market frictions. Because wage contracts are not determined at the time the schooling decision is made, a high school graduate can have less incentive to invest in college education. Had we assumed wage determination by Nash bargaining, the holdup problem might be exacerbated particularly when the workers' bargaining power is low. This would lead to a larger mismatch between returns to schooling at the aggregate level and workers' private returns. However, our model abstracts from a match function and Nash bargaining, so holdup problem is not relatively an issue.

Labor market frictions have a critical role in the wages posted. As mentioned, when firms post wages, they take into account the market conditions, the distribution of firm productivities, and the workers' schooling level. Firms are sorted with respect to their productivities over an employment spell of a worker because only high-productivity firms can post high wages. If the job-finding rate falls, there is less competition for attracting skilled workers. Therefore, firms post lower wages. If separation rate falls, then the worker has higher chances to move up on the wage ladder by accepting jobs that offer higher wages. Consequently, firms post higher wages in order to maintain the size of their labor force. Last, if job-to-job transition rate falls, low productivity firms post higher wages, whereas high productivity firms post lower wages. The intuition behind this result is that we see tighter competition among the low productivity firms because they will face a more rapid flow of workers in and out of the jobs. Also, the competition among high productivity firms is lower because they are positioned at the right end of the distribution. Hence, few workers will be able to reach that level of wages because of low job turnover rates. Whenever high productivity firms face tighter competition to attract more educated workers and post higher wages, it increases the value of college education for high school graduates and induces a higher share of college graduates in the model.

2.3.3 Equilibrium Life-time Value and Schooling Decision

In this subsection, we provide an illustrative example to discuss how an individual makes a college decision by taking into account the value of his labor market career. Figure 2 displays equilibrium supports of life-time value for skilled and unskilled worker, respectively.

Figure 2: Equilibrium Support of Life-time Values



At the steady-state equilibrium, the value of unemployment is equal to the value of employment for the marginal worker. This means that the marginal worker is indifferent between the two states. The schooling decision depends on the value of unemployment for high school and college graduates. An individual decides whether to go to college prior to labor market entry by taking into account the difference between the value of labor market career together with schooling cost. Note that the value of unemployment for an unskilled worker can be interpreted as an opportunity cost of choosing college education. Unskilled workers could move up on the wage ladder through on-the-job search during those four years that skilled workers are getting college education. This implies that the worker's opportunity cost of college education is the present value of unemployment discounted by schooling time, which is displayed in Figure 2.

The support of the lifetime value for the unskilled people is quite different from the support of the value of the skilled people. When high school graduates choose to go to college or not, they compare the discounted value of the life time value of the skilled people. After the lifetime values are discounted by the retirement rate and the interest rate together with schooling time, the support of the values almost overlap, leaving small differential in the supports between the two groups.

2.4 CALIBRATION

2.4.1 Calibration Strategy

This section calibrates the model based on the U.S. labor market data and provides the quantitative assessment of the channels that labor market friction affects for workers' schooling decision before they enter the labor market entry. In calibrating the model, we exogenously fix the interest rate and retirement rate, and then determine the other parameters, $(\delta, \lambda_U, \lambda_E, \mu, \sigma, s, \eta, b)$, using the method of simulated moments. The information in the sample that is used to define the estimator is given by M. Under the data generating process of the model, $\hat{M}(\theta)$ is assumed to be the analogous characteristics, where θ is the vector of all parameters to be calibrated. Then, the estimator is given by

$$\hat{\theta}_W = \underset{\theta \in \Theta}{\operatorname{argmin}} (M - \hat{M}(\theta))' W (M - \hat{M}(\theta)), \tag{2.25}$$

where W is a symmetric, positive-definite matrix and Θ is the parameter space. The weighting matrix, W, is a diagonal matrix and the elements of this matrix are equal to the variance of the corresponding elements of M. In order to prevent giving some parameter an extremely large weight (in the event that some parameter closely matches the value of targeted moment), all the inverse of the variance are top-coded. Then, by giving the weight W, we look for θ that makes $M - \hat{M}(\theta)$ as close to zero as possible.

Panel A in Table 2 shows exogenous parameters based on available evidence for the U.S. The annual interest rate is chosen to be five percent, which means that the discount factor r is set to be 0.0125 (Flinn and Mullins, 2015). The retirement rate is assumed to be 0.008, which implies that workers participate in the labor market for roughly 30 years on average.

Without loss of generality, the worker fixed effect on productivity for an unskilled worker s_0 is normalized as zero. Hence, we denote schooling effect by $s > 0$. The support of productivity distribution is also normalized as certain range–*i.e.*, $p \in [1, 3]$.

Table 2: Parameters

Panel A: Exogenous parameters					
Parameter	Concept	Value	Source / Criteria		
r	Rate of time preference	0.0125	5% annual interest rate		
ρ	Retirement shock	0.0080	30 years of life-time		
Panel B: Calibrated parameters					
Parameter	Concept	Value	Target	Data	Model
Schooling Decision					
μ	Schooling cost	-12.64	Tuition-income ratio	0.507	0.507
σ	Schooling Preference	12.00	College attainment rate	0.312	0.290
Worker Flow					
λ_U	Job finding rate	1.347	Unemployment spell	0.738	0.738
δ	Separation rate	0.070	Unemployment rate	0.055	0.055
Labor Compensation					
b	Unemployment benefit	0.296	Replacement rate	0.400	0.400
s	Schooling effect	0.591	College premium	1.442	1.436
η	Distribution shape	6.089	Mm-ratio	1.849	1.916
λ_E	Job turnover rate	0.683	Mm-ratio of unskilled	1.810	1.718

The schooling cost parameter μ is directly pinned down by the tuition-income ratio. The average annual wages $\overline{w}$ for a white male with a four-year college degree is 57,762 dollars (U.S. Census Bureau, Statistical Abstract of the United States: 2012, Table 232) and the average annual tuition $\overline{c}$ for four-year college is 26,021 dollars in 2010.[10] Let us denote the flow value of average annual income and schooling cost by $\tilde{w}$ and $\tilde{c}$, respectively. Then, the average annual income and schooling cost can be interpreted as follows,

$$\overline{w} = \int_0^4 \tilde{w} e^{-(r+\rho)t} dt \quad \text{and} \quad \overline{c} = \int_0^{16} \tilde{c} e^{-(r+\rho)t} dt, \tag{2.26}$$

which is paid for over four quarters. Hence, the ratio of tuition to income ($\tilde{c}/\tilde{w}$) is computed as 0.507. We target this moment by considering the ratio of flow value of schooling cost μ to average wage for college graduates.

[10] The data source is OECD (2015), Education spending (indicator). For more detailed information, see https://data.oecd.org/eduresource/education-spending.htm.

Through the equation for probability of schooling choice (2.2), the preference for schooling σ determines college attainment rate. The empirical counterpart of the college attainment rate is 31.2 percent.[11] The preference parameter σ is set to be 12.0.

The job finding rate λ_U determines quarterly unemployment duration, which is 0.738 quarters.[12] Once λ_U is set to match unemployment duration, unemployment rate for each education group u_i is determined by separation rate δ, through the equations (2.10) and (2.13). By summing up unemployment rate for each education group, we measure the aggregate unemployment rates, weighting by the steady state number of college students $s(\Delta)$ as follows:

$$u = \frac{u_0 + u_1}{1 - s(\Delta)}.$$

The empirical counterpart of the unemployment rate is 5.5 percent.

We consider the unemployment benefits b to match the replacement rate as 40 percent of minimum wages. The schooling effect s governs the college wage premium, which is here defined as the average wage ratio of college graduates to high school graduates.

As one summary statistic for wage dispersion, we use the ratio of the mean to the minimum wage (Mm-ratio: $E(w)/\underline{w}$), which is proposed by Hornstein et al. (2011). Our computation of the Mm-ratio based on the U.S. Current Population Survey March, 1980-2010 shows that Mm-ratio for high school graduates is 1.81 and 1.832 for college graduates. The overall Mm-ratio is 1.849.[13] Based on past literatures' value of this measure, which varies between 1.7 and 1.9, our values of the measure are consistent. To target these moments, we calibrate the shape parameter η of the offer distribution $H(p)$ and the job turnover rate λ_E. Given some plausible parameter values, Hornstein et al. (2011) show that most search models with a homogenous worker and firm generate a high enough reservation wage so that the observed Mm-ratio in the U.S. economy cannot be matched. However, we consider the framework in which workers face a different labor market and firms offer different wages based on workers' education level. Hence, our calibrated version of the model generates

[11] The college attainment rate is borrowed from Barro and Lee (2013). We restrict the sample only for age 25 over who completed tertiary education.

[12] The median unemployment duration is 9.6 weeks, which is equivalent to 0.738 quarters. The data source is from US. Bureau of Labor Statistics. For more detailed information, see https://research.stlouisfed.org/fred2/series/UEMPMED/.

[13] In order to reduce the measurement error, we truncate each side of the sample for five percent.

enough wage dispersion. Note that even though we do not target the Mm-ratio of a skilled worker, which is 1.832 in the data, the model generates a consistent value of the moment, which is 1.835.

2.4.2 Schooling Choice

In this subsection, we discuss how an individual makes a college education decision. When a worker with only a high school degree goes into the labor market, his life-time value is given by the weighted average of $W_0(t)$ and $U_0(t)$, where t is a period of being a worker. Unlike a worker with a high school degree, the values for college graduates get discounted by the flow value of education costs. This value is given as

$$V_t = \exp(-\phi(r+\rho))U(t) + \frac{\mu(1-\exp(-(\phi-4(t-1))(r+\rho)))}{(1-\exp(-\phi(r+\rho)))} \quad \text{for} \quad t = 1, 2, 3, 4.$$

where ϕ is a total quarter of college education and is assumed to be 16 quarters.

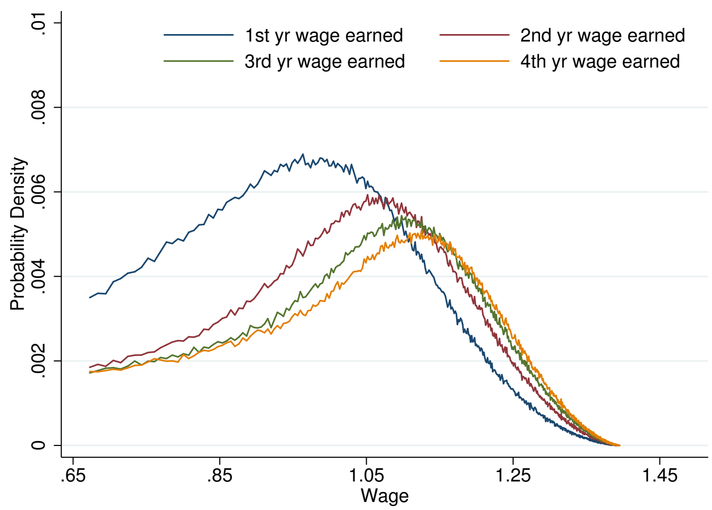
Table 3 shows how the life-time values of each education group change over time. When a worker with a high school degree enters into the labor market, an individual who chooses college education gets discounted life-time value since he will start his career after four years of education. Hence, at the starting point, high school graduates have higher values. After two years, the life-time values of college graduates become higher because it is discounted only by the remaining period of education. This shows that in the beginning, as the high school graduates' life-time value exceeds the college graduates', the opportunity cost of choosing a college education is high. However, as time goes by, college graduates catch up to high school graduates. The last column of Table 3 displays the steady state life time values for each education group.

Table 3: Changes in the life-time Value

	Life-time Values after					$\cdots$	Values at
	0 yr	1 yr	2 yrs	3 yrs	4 yrs	$\cdots$	Steady State
Skilled	36.479	43.363	50.895	59.136	68.154	$\cdots$	70.017
Unskilled	47.893	48.873	49.096	49.204	49.264	$\cdots$	49.250
Gap	-11.415	-5.510	1.799	11.724	18.890	$\cdots$	20.767

If a worker with a high school degree climbs the wage ladder while a college student gets education for a four-year education, the wage earning distribution can be a measure of the opportunity cost. This opportunity cost changes over time and depends on how fast the worker climbs the ladder, which is dependent on the labor market friction.

Figure 3: Simulated Distributions of Wage Earnings by High School Graduates



In order to investigate the changes in the opportunity cost, we simulate the model. Figure 3 shows the simulated distributions of wage earnings by high school graduates. Each distribution shows the wage earnings distribution at each year after the labor market entry. For example, if workers worked for two years right after high school graduation, then their wage earned would follow the '2nd yr wage earned' distribution in the figure. As shown in the figure, the distribution shifts to right for four years after high school graduation. This shows that workers with a high school degree typically start their labor market careers in low wage but move to better paid job through on-the-job search. When an individual makes a schooling decision, he considers how fast he can move up the wage ladder on the job without a college degree. This can be interpreted as the opportunity cost of giving up college education. If this opportunity cost is large enough, the college attainment may decrease.

In addition to impacting opportunity cost, the college attainment rate is affected by how long workers can exploit their human capital on the job. Manuelli and Seshadri (2009), Córdoba and Ripoll (2013), and Restuccia and Vandenbroucke (2014) find that adult mortality rate is one of the most important factors to explain cross country schooling differences. The intuition behind this is that changes in mortality (so as life-expectancy) heavily affect the present value of labor earnings, which is crucial for schooling choice. For a similar reason, in our paper, the duration plays a key role of schooling decision. The duration can be expressed as follows

$$ED^e = \int_0^\infty (\rho + \lambda_U) e^{-(\rho+\lambda_U)t_2} \Big(t_2 + \frac{\lambda_U}{\rho + \lambda_U} \int_0^\infty (\rho + \delta) e^{-(\rho+\delta)t_1} (t_1 + \frac{\delta}{\rho + \delta} ED^e) dt_1 \Big) dt_2.$$

At $t_2 = 0$, the total employment duration becomes

$$ED^e = \frac{1}{\rho} \frac{\lambda_U}{\rho + \delta + \lambda_U}. \tag{2.27}$$

When t_1 goes to zero, the above equation can be interpreted as the unemployment duration

$$D^u = \frac{1}{\rho} \frac{\rho + \delta}{\rho + \delta + \lambda_U}. \tag{2.28}$$

It is clear that changes in retirement shock ρ, job finding rate λ_U, and separation rate δ affect the total duration of employment and unemployment. Retirement shock ρ has a negative effect on both the employment and unemployment duration. As job finding rates λ_U increase, an unemployed searcher can easily get a job, which implies that the unemployment duration gets shorter and the employment duration gets longer. Also, as separation rate δ increases, employed workers are more likely to be unemployed, which induces longer unemployment duration and shorter employment duration.

2.4.3 Comparative Statics

We, now, discuss the implications of the model for the relationship between educational attainment and labor market search frictions. We compare the results of the U.S. baseline economy to those of the economy with a different level of search friction, keeping all else constant. Table 4 shows the results of the experiment, illustrating the effect of job finding rate λ_U, separation rate δ, and employment-to-employment transition rate on schooling.

First, we lower the job finding rate λ_U by one standard deviation. In the data, as we have discussed in Section 2.2, the college attainment decreases by 2.6 percent when the job finding rate decreases by one standard deviation. In our calibrated version of the model, compared to the result of baseline economy, college attainment decreases by 1.3 percent, when there is a one-standard deviation decrease in job finding rate. This implies that the model accounts for almost half of the effect of job finding rate on educational attainment. When the job finding rate decreases, the unemployment duration increases. This means that the period that a worker exploits the human capital shortens. Therefore, the duration effect reduces college attainment rate.

Table 4: Comparative Statics on Schooling

Decrease one	Data		Model	
stddev in	Data	% change	Model	% change
Job finding rate	0.273	-0.026	0.277	-0.013
Separation rate	0.226	-0.192	0.277	-0.011
EE Transition rate	0.335	0.196	0.287	0.024

Next, we experiment the case in which there is a one-standard deviation decrease in separation risk δ in the economy. This generates two different effects. First, an increase in the job separation rate δ induces longer unemployment duration. This implies that a worker faces lesser time that he can exploit his human capital on the job. Because of the changes in unemployment duration, the college attainment may decrease. Second, a worker with a high school degree faces higher chance to be separated from the job. That is, there is lower probability to move up on the wage ladder because of higher separation rate. This leads a worker to get more education before entering the labor market. As shown in Table 4, it seems that the latter effect dominates the former one. In the data, the college attainment decreases by roughly 19 percent in one standard deviation of separation rate, while it decreases by 1.1 percent in the model.

By decreasing the employment-to-employment transition rate by one standard deviation, the college attainment rate increases by 2.4 percent. Compared to the baseline economy, the employment-to-employment transition rate is lower, which implies that a worker with only a high school degree cannot easily move up on the wage ladder on the job while his counterpart is in college. Hence, the opportunity cost of getting a college education is lower, which increases college attainment rate.

Even though we mainly focus on the effect of labor market frictions on educational attainment, it is worthwhile to look at the effect of the retirement shock ρ. With the available data, we arbitrarily change the retirement shock. When we reduce the total employment period by five years, the college attainment rate decreases by about 8.4 percent compared to the baseline economy. This is because a decrease in ρ extends the period that a worker exploits his human capital on the job. More specifically, as their expected unemployment duration is shorter, they are more likely to choose higher education rather than enter the labor market earlier.

2.5 EFFICIENCY ANALYSIS

In this section, we discuss the efficiency properties regarding college attainment in the model. We, following Sim (2015), consider the problem of the constrained planner who maximizes the present value of the expected output flow by a newly born worker throughout her life. An unskilled worker is expected to produce $p+s_0$ when she works at a firm with productivity p, whereas a skilled worker is to produce $p+s_1$. The present value of the expected output flow by an unskilled and skilled worker working at a firm with productivity p is denoted by

$S_{0e}(p)$ and $S_{1e}(p)$, respectively. For each $i \in \{0, 1\}$,

$$rS_{ie}(p) = p + s_i - \rho S_{ie}(p) + \delta(S_{iu} - S_{ie}(p)) + \lambda_E \int_p^{\overline{p}} [S_{ie}(p') - S_{ie}(p)] dH(p'), \tag{2.29}$$

where S_{iu} represents the present value of the expected output flow by an i-type unemployed worker. Regardless of their education, both college and non-college graduate workers produce b through non-market activity when they are unemployed, which is given by

$$rS_{iu} = b - \rho S_{iu} + \lambda_U \int_{\underline{p}}^{\overline{p}} [S_{ie}(p') - S_{iu}] dH(p'). \tag{2.30}$$

Then, the social planner makes schooling decisions on behalf of individual workers as follows. Given idiosyncratic preference ε, the planner sends the worker to college if and only if

$$\exp(-(r + \rho)\Delta) S_{1u} + \varepsilon \geq S_{0u}. \tag{2.31}$$

Applying 'integration by part' to (2.29) and combining it with (2.30) yields

$$\Big[r + \rho + \delta - \frac{\lambda_U \delta}{r + \rho + \delta}\Big] S_{iu} = b + \frac{\lambda_U}{r + \rho + \delta} \Big[\frac{1}{2}(\overline{p}^2 - \underline{p}^2) + s_i(\overline{p} - \underline{p}) + \int_{\underline{p}}^{\overline{p}} \lambda_E \mathbb{E}[S_e(p') | p' \geq p] dp\Big], \tag{2.32}$$

where

$$\mathbb{E}[S_e(p') | p' \geq p] = \int_p^{\overline{p}} \frac{1 - H(p')}{r + \rho + \delta + \lambda_E(1 - H(p'))} dp', \tag{2.33}$$

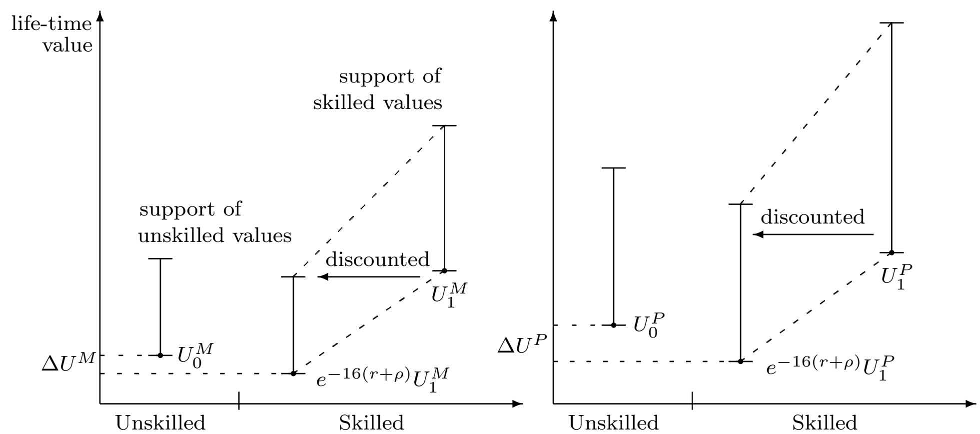
and

$$\begin{aligned} \frac{d(S_{0u} - \exp(-(r + \rho)\Delta) S_{1u})}{d\lambda_E} &= [1 - \exp(-(r + \rho)\Delta)] \mathbb{E}[S_e(p') | p' \geq p] \\ &\quad \times \frac{\lambda_U}{r + \rho + \delta} \Big[r + \rho + \delta - \frac{\lambda_U \delta}{r + \rho + \delta}\Big]^{-1}. \end{aligned} \tag{2.34}$$

In particular, as $\Delta \to 0$, the planner's schooling decision becomes independent of λ_E. While the planner's schooling decision is not directly affected by the job turnover rate, individual's schooling decision in the market equilibrium is negatively associated with the job turnover rate. It suggests the possibility of under- and over-education from an efficiency perspective.

Figure 4 displays the opportunity cost of college education at the high school graduation date. When an individual makes a college education decision, he may compare the expected value of labor market career as a high school and college graduate. Panels (a) and (b) in Figure 4 clearly show that there is over-investment in schooling compared to the social planner's problem if we compare the discounted life-time values of high school and college graduates–i.e., $\Delta U^M > \Delta U^P$.

Figure 4: The Opportunity Cost of Schooling at High School Graduation



(a) Market Equilibrium (b) Planner's Problem

The previous literature, such as Charlot and Decreuse (2005), Charlot and Decreuse (2010), and Flinn and Mullins (2015), also discuss about the market efficiency in education investment. Charlot and Decreuse (2005) and Charlot and Decreuse (2010) develop a matching model with educational choice and show that there might exist over-investment in education under search frictions. In their set-up, the firm's conditional expectation on ability of workers is affected by the schooling decisions of workers. That is, more workers choosing college education becomes negative externality for the market. Since the social planner does not internalize the negative externality, the social planner forces workers with low ability to not choose college education. Hence, compared to the social planner problem, the market equilibrium generates over-investment in education. Even though Flinn and Mullins (2015) use a similar model, they predict a different result. This is because their results highly depend on the surplus sharing rule. The under-investment in education in Flinn and Mullins (2015) result from weakness in the workers' bargaining position. One thing worth noticing is that Charlot and Decreuse (2005), Charlot and Decreuse (2010) and Flinn and Mullins (2015) overlook the opportunity cost during the schooling period by assuming no time cost for schooling.

As mentioned in the previous section, the over-investment in college education problem might be resolved by increasing the opportunity cost of entering the labor market as high school graduates. That is, if workers with only a high school degree have more stable jobs, then the college attainment would decrease. The second column of Table 5 shows the market efficiency of educational investment when there is higher opportunity cost compared to the benchmark economy. Even though the college attainment rate decreases by 3.7 percent, the rate generated by the social planner also decreases by almost same amount. Hence, in this case, the college attainment would still be high enough to raise the over-investment problem, unlike the problem of college attainment under a social planner.

Improving the quality of college education is another way to improve educational attainment's social efficiency. As the quality increases, the gap between the market equilibrium level of college attainment and social efficient level gets narrrower. The third column of Table 5 shows that under our parameterizations if the quality of college education increases by 38 percent in terms of productivity, then the market equilibrium college attainment achieves the social efficient level.

Table 5: Social Efficiency in Education Investment

	B.E.	Increase in job security of unskilled	Increase in quality of education
ME	0.296	-	+
PP	0.278	-	+
Ratio	1.065	+	-

2.6 CONCLUSION

This paper analyzes the channel through which labor market friction affects individuals' schooling decision prior to labor market entry. It demonstrates that a country with a higher job turnover rate and/or lower job separation rate tends to incur a higher opportunity cost of schooling and maintain a lower college attainment rate. The calibrated model reconciling the U.S. data predicts that the market equilibrium results in over-investment in schooling from an efficiency perspective. The opportunity cost of schooling (compared to the expected gains from it) at the time of high school graduation is smaller in the market equilibrium than the planner's problem.

Naturally, the next step in this research is to study how experience on the job and on-the-job training affect the choice of college attainment. We conjecture that if workers with only a high school degree could get on-the-job training or work experience, then they would climb the wage ladder faster for the first four years. As a result, the opportunity cost of college education would increase, lowering college attainment. We leave this topic for future research.

3.0 POLICY-INDUCED VARIATION IN COLLEGE LABOR SUPPLY AND THE SKILL PREMIUM[1]

3.1 INTRODUCTION

Two well-documented educational trends have affected most advanced economies. First, there is a broad increase in educational attainment (Restuccia and Vandenbroucke, 2013b). Second, there has been a general increase in the educational wage premium, whether measured as the estimated return to an additional year of schooling or as the college-high school wage premium (Machin and Van Reenen, 1998; Bekman et al., 1998; Katz and David, 1999). The goal of this paper is to explore whether these two trends are linked through composition effects.[2] The idea is that the expansion of college is likely to have changed the average ability of high school and college graduates, where ability is understood broadly as unobserved characteristics that affect wages. We want to know whether these changes may have contributed to the rise in college wage premiums.

An existing literature has sought to explore this question with a primary focus on the United States. Here, we revisit these questions in the context of Korea. Korea experienced an educational expansion and wages changes similar to those in other advanced economies, as can be seen in Figure 5.[3] Figure 5a shows the educational attainment of Koreans by birth cohort, with attainment measured in four complete and mutually exclusive categories: those with less than a high school degree; high school graduates; those with more than a high

[1]This research is a joint work with Jaehan Cho and Todd Shoellman

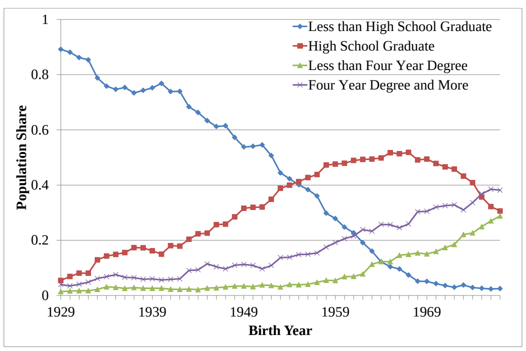
[2]By contrast, a well-known alternative literature explores a link through relative price effects: if workers with different education levels are imperfect substitutes, then changes in relative supply necessarily affect relative wages. However, this force generally tends to imply a decline in the college wage premium, all other things equal, and so is not helpful in understand recent experiences (Katz and Murphy, 1992).

[3]Details on the construction of this and subsequent data are delayed to Section 3.2.

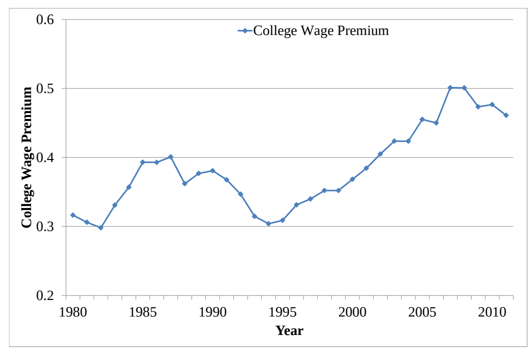
school degree but less than a four-year college degree; and four-year college graduates. The expansion of education is clear: while 90 percent of the 1930 cohort did not graduate high school, today more than half of the Korean population attends college. Figure 5b shows the college wage premium by year from 1980 onward. The wage premium fluctuated until 1995 but has risen 15–20 log points since that time.

Figure 5: Changes in Korean Education

(a) Expansion of Education



(b) College Wage Premium



While the broad patterns in Korea are similar to those of other advanced economies, Korea does offer one special advantage: the expansion of education was unusually large and rapid. For example, about thirty years separate the first birth cohort to achieve 10 percent college graduation and the first to achieve 30 percent college graduation in the United States; the same figure is about ten years in Korea. And while the college graduation rate has leveled off at about 30 percent in the United States it has risen to roughly 40 percent in Korea.

To exploit this rapid expansion of education we adopt two different empirical strategies previously used in the literature to quantify various channels through which educational attainment might affect wages. While the exact identifying assumptions vary by strategy, they all share in common the idea of comparing educational wage premiums across cohorts with different educational attainment at a given point in time. Korea offers a clear advantage for such empirical approaches because of the larger differences in educational attainment for nearby cohorts, which we exploit.

We then go a step further than the existing literature by exploiting a policy-induced expansion of educational attainment. The Korean government limited university enrollment throughout our period of interest. By the late 1970s the limit was sufficiently binding that it encouraged the growth of a large tutoring industry that help high school students score better on the college entrance exam. The change in government after the assassination of President Park in 1979 brought large policy changes throughout the economy. Of particular interest to us is a large and sudden increase in university enrollments in 1981 and 1982. Our regression discontinuity analysis below suggests that this policy increased the post-policy university attainment by about 3 percentage points (about 15 percent) as compared to pre-policy trends. We exploit this exogenous increase in enrollment within the framework of the existing empirical strategies for alternative, exogenous variation in the quantity of schooling.

Our analysis yields two main results. The first concerns the importance of composition effects in accounting for the cross-sectional return to schooling. In line with Kaymak (2009), we find that composition effects account for nearly one-half of the observed return to schooling, indicating that the true private return to schooling is slightly more than half of the observed Mincer return. The magnitude of this effect is consistent regardless of whether we use all educational variation or only exogenous educational variation to estimate the effect. The second result concerns the importance of composition effects in the time series. Consistent with Juhn et al. (2005), we find evidence that increasing college attainment lowers the college wage premium for a cohort, suggesting that higher college attainment is obtained by lowering the relative ability of college graduates. However, the results here depend somewhat on whether we use all the variation or only policy-induced variation in educational attainment. The former suggests a small and relatively weak effect, while the latter suggests a much stronger effect.

In addition to the papers listed above, we are also closely related to a number of other studies that investigate similar issues, primarily within the United States. Laitner (2000) formulates a model that generates qualitative predictions in line with what we study here, but does not attempt to quantify these forces. Recently Bowlus and Robinson (2012) use the flat spot method of Heckman et al. (1998b) to try to identify changes in skill prices versus skill quantities in explaining changes in wages for four educational groups in the U.S.; changes in innate ability can be thought of as one source of changes in skill quantities. Carneiro and Lee (2011) use a different empirical strategy that relies on controlling for all possible sources of skill price variation to help identify skill quantities as a residual; again, ability is one component of the quantity of skill. Carneiro and Lee (2011) use a local instrumental variable approach to predict the wage implications of expanding college enrollment. Unfortunately it is difficult to provide a consensus result from these papers because the literature has yet to reach one. Some papers find modest composition effects (Juhn et al., 2005; Carneiro and Lee, 2011), but others find sizable ones (Kaymak, 2009; Bowlus and Robinson, 2012). Our hope is to provide further evidence to this debate by exploiting the advantages of the Korean experience outlined above.

The rest of the paper proceeds as follows. Section 3.2 reviews the Korean data and the relevant details about Korean educational policy. Section 3.3 conducts the analysis. Section 4 concludes.

3.2 EDUCATION AND EDUCATION POLICY IN KOREA

In this section we outline briefly the relevant details of the Korean educational experience. Our focus is on two main aspects. First, we highlight the post-World War II trends, which include a large expansion of education and a recent increase in the college wage premium. Second, we highlight the role of exogenous policy changes in the 1980s in affecting the educational expansion.

We measure educational attainment using the Korean population censuses conducted in 1966 and then every five years from 1970 to 2010. Throughout this period the census has contained a question on the highest educational attainment of respondents.[4] We code respondents into four broad categories that are comparable over time: those with less than a high school degree; high school graduates; those with more than a high school degree but

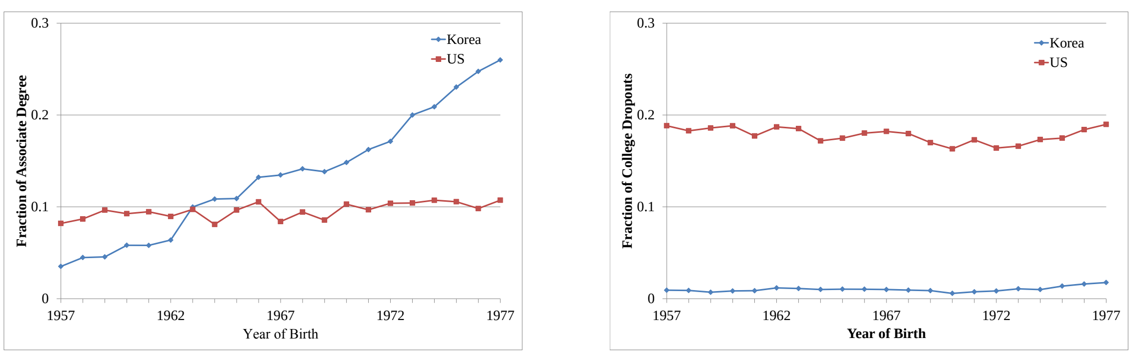
[4]This is also a minor advantage as compared to the United States. There the main data sources are the Current Population Survey and the Population Census. They asked only about years of schooling and not attainment until 1992 and 1990, which generally forces researchers to assume that, say, workers with 12 years of schooling are high school graduates and so on.

less than a four-year college degree; and those with at least a four-year college degree. We measure attainment for the 1929–1977 cohorts using the census taken when they were aged 33–37. Using this five year window with censuses taken every five years gives us exactly one observation of attainment per cohort. The results are plotted in Figure 5a. Korea has experienced a larger and more rapid rise in education than the United States. We show in the appendix B.3 that similar patterns hold for men and women separately.

Examination of Figure 1 shows that a substantial fraction of Koreans have more than a high school degree but less than a four-year college degree. A similarly large group is present in U.S. data (Hendricks and Schoellman, 2014). However, this group contains two subgroups: those who obtain a two-year degree and those who start college but obtain no degree. Further, the relative proportion of the subgroups varies greatly between the U.S. and Korea. This is shown in Figure 6. Roughly two-thirds of the "some college" group in the U.S. consists of those who obtain no degree, whereas in Korea almost all students with less than a four-year degree obtain a two-year degree.[5]

Figure 6: Decomposing the Some College Group in the U.S. and Korea

(a) Two-Year Degrees (b) No Degree



A final advantage of using Korean data is that variation in educational policy provides exogenous variation in educational attainment. The most important change for our purposes came in the summer of 1980. High school education was available to all students and paid for

[5]U.S. data taken from the March Current Population Survey.

by the federal government since 1968. However, the federal government strictly controlled university admissions. Students took two examinations to determine who would be admitted (Lee, 1992). As the fraction of students who graduated high school rose, the restriction on enrollment became more binding. By 1980, nearly half of all high school students were enrolled in after-school private tutoring to help improve their college enrollment test scores.

Education policy changed discretely after the 1979 assassination of President Park. General Cheon assumed control of the country in 1980 and was recognized as President in 1981. He instituted a host of reforms throughout the economy. The educational reform had two key components. First, he banned private tutoring. Second, he greatly expanded educational enrollments through several mechanisms: by opening new universities; by expanding the departments per university; and by expanding the students per department. The aggregate effect was large: total new enrollments were nearly 50 percent larger in 1981 as compared to 1980, and enrollment in four-year degrees were 60 percent larger. This large, exogenous increase is critical to our empirical approach.

In order to utilize the effect of this policy in our empirical work, we need to define which cohorts were affected by it. It is clear that all cohorts who were age 19 or younger in 1981 (born 1962 or later) was affected by the policy. Those who were slightly older may have also benefited from a second chance to enter university. This is less clear in the case of Korea, because many who were marginally denied admission to four-year colleges would have been accepted to and attended two-year colleges instead. Roughly two-thirds of applications to college in the early 1980s were from high school seniors taking the college entrance for the first time; the remainder were from repeat test-takers.[6] Nonetheless, we think of those born 1962–1964 as having potentially been marginally affected by the policy. Those born before 1962 were likely not affected. Hence, we intend to exploit the large rise in college attendance and completion between cohorts born before this time and those born just after the policy came into effect.

Figure 7 isolates the fraction of each birth cohort that obtains at least a four-year college degree. The vertical line marks the 1962 cohort, which was the first to be fully affected by the policy. By comparing the attainment of cohorts on either side we can see that the

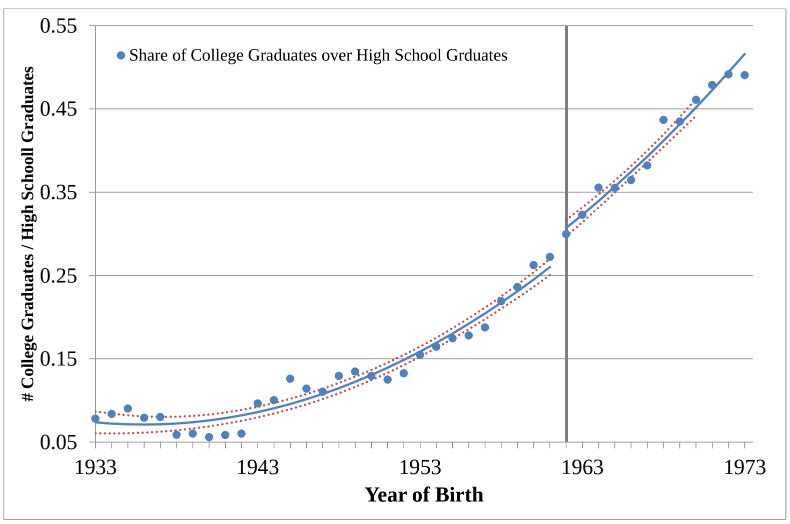
[6]Data from the Korea Statistical Yearbook, 1985.

policy did indeed increase attainment. To make this statement more precise we estimate the effect of the policy using a regression discontinuity approach. We consider the following specification:

$$S_{ic} = \alpha + \tau D_i + \beta g(c) + \varepsilon_i \tag{3.1}$$

where the dependent variable, S, is an indicator variable that takes on value 1 if individual i from cohort c obtained a four-year degree. We predict this value under the assumption that there is a smooth trend relationship between attainment and cohort captured by $g(c)$. D is an indicator which takes the value of 1 if $c \geq 1962$ and 0 otherwise. We are interested in τ, the estimated jump effect of the policy on educational attainment.

Figure 7: Fraction College Graduates by Birth Cohort



Note: Figure shows college graduation rates by birth cohort, taken from the Korean census. The vertical line indicates the birth cohort that was marginally exposed to the educational reform; younger cohorts to the right were fully exposed, while older cohorts to the left were not exposed at all.

Table 6 gives the estimated results when we assume that the underlying trends in cohort $g(c)$ are captured by quadratic polynomials. We estimate a discontinuous effect of the policy on college attainment of 3.2 percentage points (column 1), or 2.1 percentage points if we focus only on the fraction of students who graduate at least high school (column 3). One slight complication is that we are using educational attainment data for people drawn from different ages. We find that controlling of this possible confounding effect using a quartic polynomial in age makes no difference for our results (columns 2 and 4). Thus, policy induced a large change in college attainment in Korea. To put the figure into context, note that total college attainment for the 1980 cohort was around 25 percent, and that the policy induced an additional (above trend) increase of 3.2 percentage points, or roughly 12 percent of total 1980 attainment. This effect is large and gives us hope to provide alternative identification of composition effects.

Table 6: Estimated Increase in College Attainment Induced by Policy Change

	(1)	(2)	(3)	(4)
dependent variable:	Total Fraction		Fraction of High School Graduates	
	0.032***	0.033***	0.021***	0.022***
	[0.009]	[0.008]	[0.007]	[0.007]
Birth Cohort Polynomial	Quadratic	Quadratic	Quadratic	Quadratic
Age Polynomial Controls	No	Quartic	No	Quartic
Initial Sample Size	253896	253896	160204	160204

Note: The dependent variables are whether the individual graduates a college. Each coefficient is from a separate regression. Each regression includes controls for a birth cohort quadratic polynomial and an indicator whether or not a cohort entered a college after the educational reform. The bracketed values indicate corresponding standard errors.

Finally, we turn our attention to the evolution of the college wage premium in Korea. The Korean census does not collect data on wages. Instead, we use the Korean Survey Report on Wage Structure's annual data from 1980–2011. The survey collects data about the characteristics of workers and their compensation from firms with ten or more regular workers.[7] The important data for our purposes are each worker's final education degree, their age and gender, and their labor market earnings. The survey is large, containing roughly half a million workers each year.[8]

[7]Since 1998, the sampling criteria has been extended to firms with 5 or more regular workers. We use the information on firms with 10 or more workers only to maintain comparability of the data over time.

[8]Further information on this survey can be found at http://laborstat.molab.go.kr

In order to measure the skill premium we construct a sample along the lines of Katz and Murphy (1992). We include only full-time, full-year workers aged 18–65 who worked at least 35 hours per week at the time of the survey. We define hourly wages using monthly income and hours worked per month. We use the CPI to deflate all wages to 2010 dollars. Individuals whose real wages are less than \$4.11 per hour (the 2010 minimum wage) are excluded from the sample.

We estimate the college wage premium by regressing log-hourly wages on dummies for educational attainment, controlling for age, gender, and potential experience interacted with gender, where potential experience is defined as age minus years of schooling minus 6. The college wage premium is the estimated coefficient for having graduated with a four-year college degree minus the estimated coefficient for having only a high school diploma. Figure 5b shows the results. The college wage premium fluctuated between 0.3 and 0.4 between 1980 and 1995; from 1995 to 2010 it rose from about 0.3 to 0.45, with some modest signs of falling recently. In the next section we explore whether the changes in the college wage premium can be linked to the changes in educational attainment.

3.3 EMPIRICAL FRAMEWORK

We now turn to the empirical analysis. Our goal here is to consider several approaches proposed in the literature. As we will see, these approaches generally rely on comparing nearby birth cohorts with different education levels. Our goal is to exploit two features of the Korean data. First, the rise in education was much sharper. Since identification rests on comparing across cohorts, this is an advantage. Second, we have exogenous, policy-induced variation in the supply of education.

It is useful to provide a unified treatment of the problem. To do so, we focus on the following wage equation:

$$\log(w_{itc}) = \beta_t + \beta_c + \beta_{t-c} + (\gamma_t + \gamma_c + \gamma_{t-c})S_i + a_i + \varepsilon_i, \tag{3.2}$$

where w denotes the hourly wage, S denotes schooling, a denotes ability, and ε is the error term. At some points we focus on the college-high school wage premium, in which case $S_i \in \{0, 1\}$ will denote high school and college graduates, respectively; at other we focus on the Mincerian return to schooling, in which case $S \in [0, \bar{S}]$ will be a continuous variable. Throughout, we use subscript i for individuals, t for time, and c for birth cohort. It follows that $t - c$ is age. We use β_t as a shorthand for a full set of year dummies, and similarly for the remaining βs and γs. Note that we allow both the level of wages and the return to schooling to depend in an arbitrary way on age, year, and cohort. It is well-known that at this level of generality these effects are not well-identified because of a linear dependence among the three; we return to this point further below.

Our goal is to understand the role that ability plays in the patterns of average wage by school group, given by:

$$E[\log(w_{itc})|S] = \beta_t + \beta_c + \beta_{t-c} + (\gamma_t + \gamma_c + \gamma_{t-c})S + E[a_i|S]. \tag{3.3}$$

We want to understand the importance of composition effects for wages, by which we mean differences in $E[a_i|S]$ across cohorts and time. We will explore both discretized and continuous schooling models, in which case the relevant expressions are

$$E[\log(w_{itc})|1] - E[\log(w_{itc})|0] = \gamma_t + \gamma_c + \gamma_{t-c} + E[a_i|1] - E[a_i|0] \tag{3.4}$$

and

$$\frac{\partial E[\log(w_{itc})|S]}{\partial S} = \gamma_t + \gamma_c + \gamma_{t-c} + \frac{\partial E[a_i|S]}{\partial S} \tag{3.5}$$

Composition effects affect wages unless $E[a_i|1] - E[a_i|0] = 0$ or $\frac{\partial E[a_i|S]}{\partial S} = 0$. However, quantifying composition effects is generally challenging, which likely explains the diversity of approaches and results in the literature. There are two main obstacles. First, ability is not observed directly. Some datasets include proxies for ability (such as standardized test scores), but it is necessary to account for the noise inherent in test scores (Taubman and Wales, 1972; Bishop, 1989; Hendricks and Schoellman, 2014). The Korean data, and most other large datasets worldwide such as censuses, lack such proxies. A second challenge is that both of these wage equations suffer from a classic collinearity between age, time, and cohort effects. This collinearity is worsened by the inclusion of ability conditional on schooling. Since schooling is usually fixed by cohort in the empirical analysis, this implies that mean ability conditional on schooling is itself another cohort effect.

In some cases it is possible to rule out the importance of age, time, or cohort effects, but that does not appear to be the case here. It is well-known that the college wage premium varies by age. Time effects are naturally suggested by the typical framework that models the wages of college and high-school educated workers as a function of the quantity of workers with the two types of skill and the prevailing level of skill-biased technical change in the economy (Katz and Murphy, 1992; Goldin and Katz, 2008). We have already discussed that cohort effects capture (at least) the role for differences in ability conditional on schooling.

Thus we conclude that empirical progress depends on confronting these two challenges. We now discuss and implement the approaches suggested in the literature. Throughout, we emphasize how composition effects can be identified and separated from time or age effects.

3.3.1 Composition Effects in the Cross-Section

Our first approach follows Kaymak (2009). Kaymak works with the continuous school model in equation (3.2) and a linear return to schooling. He uses a cohort-based instrumental variable approach to estimate the return to schooling. Intuitively, a valid instrument allows one to measure the true return to schooling $\gamma_t + \gamma_c + \gamma_{t-c}$, whereas OLS estimates are biased and yield $\gamma_t + \gamma_c + \gamma_{t-c} + \frac{\partial E[a_i|S]}{\partial S}$. Simple subtraction yields an estimate for the cross-sectional role of ability bias $\frac{\partial E[a_i|S]}{\partial S}$. Note that this approach also sidesteps the age-time-cohort problem by not attempting to disentangle them at all.

Kaymak proposes using cohort dummies as instruments. Cohort dummies are clearly exogenous to an individual and highly correlated with school attainment (even more so in Korea). It is less clear whether they satisfy the exclusion restriction. That restriction requires that cohort only affects wages through its affect on average educational attainment. It precludes effects that might arise through, say, labor market conditions upon first arrival to the labor market, which would likely be common to a cohort. Kaymak (2009) proposes and implements a number of controls to help capture such effects. Here we explore an alternative approach, which is to use our educational policy change as an alternative instrument that affected schooling and satisfies the exclusion restriction.

We start by following Kaymak (2009) closely. We estimate the effect of years of schooling on real wage using the data from the Korean Survey Report on Wage Structure.[9] We implement our regression as:

$$\log(w_{itc}) = \beta_t + \gamma S_i + \beta X_i + \varepsilon_i.$$

where X includes control variables: a quartic trend in cohort (to capture slow-moving, cohort-specific trends such as changes in mean ability (Flynn, 1984, 2007)); a quartic function of age (to capture returns to experience); and dummies for survey year. We estimate this equation three ways: by OLS, and then using two different instruments.

Table 7: Return to Education

	Dependent Variable : log(*Real Wage per Hours*)		
	LS	IV	
	(1)	(2)	(3)
100 × Years of Education	7.01***	4.34***	2.85***
	[0.006]	[0.143]	[0.372]
Instruments	N/A	Year of Birth	Educational Reform
Observations	8,018,485		
R-squared	0.480	0.392	0.392

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Note: We include a quartic age trend, a quartic year of birth trend and dummies for survey year. The instrument in column (2) is a full set of birth cohort dummies; in column (3) it is a dummy for cohorts born in the three-year window 1961–1963 who were most suddenly affected by the educational reform.

Table 7 displays the estimation results. The estimates in columns (1) and (2) are quite comparable to Kaymak (2009) and indicate that a little more than half of the return to education (4.34/7.01) is true return to schooling, while the other half is due to composition effects. In column (3) we explore our alternative instrument. Our instrument variable is dummy for the cohorts who were born between 1962–1964. We find out that a little more

[9] The detailed description of the sample is provided in Appendix B.

than one over third of the return to education (2.85/7.01) is true return to schooling and composition effects are a little larger than the result in column (2).

We then follow Kaymak (2009) by estimating the same equation separately for each decade. Our goal here is to assess whether the patterns of the skill premium observed in Figure 5b are accounted for by composition effects or by changes in skill prices. The results are given in Table 8. The least squares estimates are in line with the changes in the college wage premium plotted in Figure 8, showing first a decline in the 1990s and then a pronounced rise to higher levels in the 2000s. In column (2) we again show the result of instrumenting for attainment using cohort as in Kaymak. We find evidence that the return to schooling is increasing over time. Taking the difference between columns (1) and (2) suggests a nonlinear pattern: ability bias was the largest in the 1980s, declined in the 1990s, and rose again in the 2000s. In column (3) we again explore the role of using our alternative instrument that controls 1962–1964 born cohorts. As we already show for the whole period, true return to schooling is smaller than when we use the instrument in Kaymak. When we use our instrument variable, true return to schooling is statistically zero in 1980s and 2000s.

Table 8: Patterns of the Return to Education: 1980–2010

	LS	IV	
Year	(1)	(2)	(3)
1981-1990	7.2***	3.8***	1.6
	[0.008]	[0.419]	[1.495]
1991-2000	5.8***	4.1***	2.9***
	[0.009]	[0.261]	[0.497]
2001-2010	8.7***	1.9***	-0.7
	[0.014]	[0.269]	[0.611]

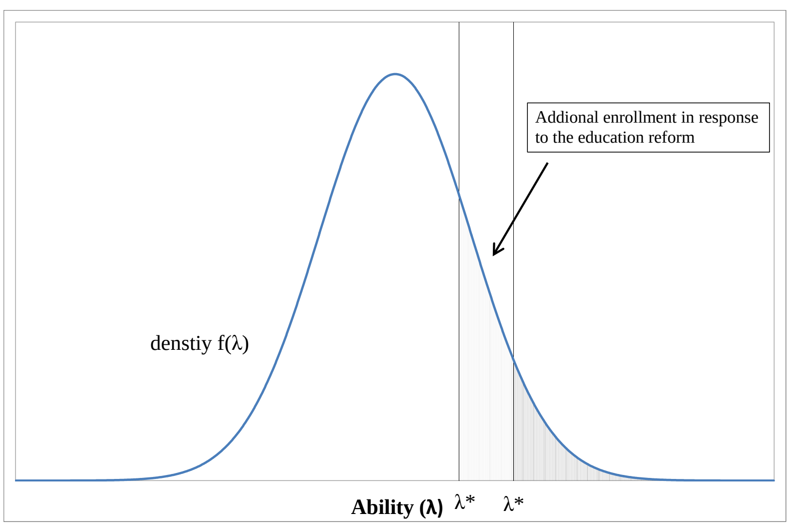
Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Notes: We control for a quadratic age trend, a quadratic trend in year of birth, survey year, cohort size by education and survey year

3.3.2 Composition Effects in the Time Series

Our second approach follows Juhn et al. (2005). These authors work with the discretized school model in equation (3.4) and discrete schooling groups. Their approach is based on an assumption about how cohort effects will enter the wage equation. The intuition is conveyed by Figure 8, which draws closely on a similar figure in their paper. To simplify, suppose that ability is (log-) normally distributed and that selection by ability into educational attainment is perfect. In this case there is an ability cutoff for each cohort λ^* such that all higher-ability individuals complete college and work as college graduates, while all lower-ability individuals do not attend college and work as high school graduates. An expansion of college attainment means a shift left in the cutoff, so that the individuals that are shaded light gray now graduate college instead of only completing high school. Under this simple and clear model of selection effects, mean relative ability by cohort $E(a_{i,t-v}|s=1) - E(a_{i,t-v}|s=0)$ can be proxied for by using the fraction of the population that attains college for each cohort.

Figure 8: A Simplified View of Composition Effects



This approach obviously embeds strong assumptions about how ability effects work. Hendricks and Schoellman (2014) find that there are two important complications to this view of selection effects. First, the sorting of students ability into attainment is not perfect, as the figure suggests. Second, there are not only two mutually exclusive school categories as that figure suggests. Hendricks and Schoellman (2014) find that both of these factors play an important role in the U.S. Their results suggest, for example, that the expansion of education had essentially no effect on the mean ability of college graduates, although it did cause a decline in the mean ability of high school graduates. These findings suggest it may be worthwhile to explore alternative assumptions.

We do so by once again exploiting our exogenous educational policy change. These approaches exploit the idea that while long-term changes in educational attainment may be complicated by simultaneous trends in sorting and the attainment of some college, large short-run changes are unlikely to be. Hence, instead of asking whether the growth in college attainment over the long run has caused changes in the college wage premium, we ask whether the large growth in college attainment induced by the policy reform cause changes in the affected cohorts' college wage premium.

We start by following Juhn et al. (2005) closely. They suggest the equivalent of differencing equation (3.4) over time,

$$\Delta \frac{E[\log(w_{itc})|s=1]}{E[\log(w_{itc})|s=0]} = \Delta\gamma_c + \Delta\gamma_{t-c} + \Delta[E(a_i|1) - E(a_i|0)].$$

which has the effect of netting out time effects. They assume that cohort effects and changes in mean ability are accounted for by changes in average educational attainment, $\Delta\gamma_c + \Delta[E(a_i|1) - E(a_i|0)] = \omega\Delta[E(s_{itc})]$ We use a regression of education on age and cohort dummies to predict for each cohort the average attainment at age 35, which we use as $E[s_{itc}]$.[10]

Before giving the regression results it is useful to look at the raw relationship between the college wage premium and the fraction of the population with a college degree. The two are plotted together in Figure 9. Each panel of this figure corresponds to one five-year differenced comparison; for example, the top left panel shows the difference in wages and attainment when comparing similarly aged workers between 1980 and 1985. The blue line in each figure shows the five-year change in the college wage premium, while the red line

[10] The detailed description of the sample is provided in Appendix B.2.

shows the five-year change in the fraction of the population who graduates college. The correlation between the two is generally weak, much weaker than what Juhn et al. (2005) find in the United States. However we also shade in each figure in gray the cohorts that were marginally affected by the educational reform. We can see in each figure that the reform consistently induced an unusually large increase in attainment and that the college wage premium consistently declined.

To check these visual results we estimate the above equation using weighted least squares. Table 9 reports the estimation results. In column (1) we follow Juhn et al. (2005) and estimate the change in the college wage premium as a function of the change in the share of the cohort in college. We find a negative result consistent with their work, but the estimate is imprecise. In columns (2)–(5) we explore the value of incorporating our educational reform. In these regressions we continue to include the share of the cohort that graduates college, but we also include a dummy variable that takes the value of 1 for cohorts that were marginally affected by the educational reform. The difference between columns (2)–(5) lies only in the which cohorts we define as having been affected. In column (2) we use a strict definition, the 1962–1964 cohorts. In column (3) we expand the affected cohorts to the 1962–1966 cohorts; in column (4) we expand it to the 1962–1971 cohorts; and last in column (5) we expand it to the 1962–1976 cohorts.

The results of regressions (2)–(5) agree closely. In each case the effect of cohort college share is of the wrong sign and is statistically insignificant. This indicates that for cohorts not affected by the educational reform, there is no strong relationship between the fraction graduating college and college wage premiums. As explained above this could be the case if offsetting forces (such as changes in the structure of education, education quality, or the sorting of students) offset the increase in attainment. Indeed, it could be the case that such forces are themselves responsible in part for causing the increase one educational attainment. The regressions also agree closely on the effect of the educational reform. Marginally affected cohorts have statistically and economically lower college wage premiums, on the order of 18–23 log points, as compared to the unaffected cohorts (those born earlier or much later). As we expand the scope of the number of cohorts affected, we find that the estimated effect declines. The intuition is that the rapid expansion in education that took place at this time

Figure 9: Log Change in Relative Wage and College Share

was exogenous and happened too quickly to be offset by other factors. In this case, the policy had the expected effect consistent with the findings of Juhn et al. (2005).

Our preferred interpretation of this finding is that long-run trend changes in educational attainment may come in part from better sorting by ability or changes in educational preparation, in line with Hendricks and Schoellman (2014), while sudden short-run changes in educational attainment cannot and hence cause stronger decreases in ability and wages.

Table 9: Effects of Cohort-Specific College Share on Wages of College Graduate Men

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable : $\Delta \frac{E[\log(w_{itc})\|s=1]}{E[\log(w_{itc})\|s=0]}$				
Cohort College Share	-0.081	-0.009	0.065	-0.011	-0.050
	[0.130]	[0.134]	[0.136]	[0.134]	[0.135]
Cohort Size	-0.249**	-0.243**	-0.275***	-0.282***	-0.269**
	[0.103]	[0.103]	[0.101]	[0.104]	[0.106]
Education Reform		-0.188**	-0.233***	-0.129*	-0.057
		[0.095]	[0.077]	[0.066]	[0.069]
Constant	-0.166	-0.192	-0.211	0.214	0.419***
	[0.154]	[0.153]	[0.152]	[0.152]	[0.153]
Observations	230	230	230	230	230
R-squared	0.321	0.335	0.353	0.335	0.324

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Notes: We use wage information for the dependent variable from Report on Wage Structure Survey. In each column we use predicted college share of each cohort at age 35 as specified in equation (1) in Juhn et al. (2005). The regression also includes controls for cohort size, seven year dummy variables, and two age dummy variables for young and middle age workers. Education Reform is a dummy that indicates cohorts affected by the educational reform as explained in the text. Each column describe the cohorts who were born within 3 years in (2), 5 years in (3), 10 years in (4) and 15 years in (5) after the education reform passed.

3.4 CONCLUSION

In this paper, we study the importance of composition effects for education wage premiums in Korea. Korea offers an interesting case because it has experienced a rise in educational attainment and education wage premiums, like most other advanced economies. However, the rise in education was much larger and more rapid than that in most other advanced economies. Further, some of the increase was policy-induced.

We use a simple empirical framework to introduce the main challenges confronted in the literature. We then implement several procedures suggested in the literature to overcome these challenges. We also show how to modify the existing approaches to exploit policy-related variation in educational attainment.

Our findings are robust across the approaches and specifications. We find only modest evidence for composition effects in the long run. By this, we mean that more-educated workers are of higher average ability, but there is little change in the strength of this relationship over time. However, we find strong evidence for a decrease in education wage premiums for the cohorts most affected by the policy-induced expansion of education. This suggests to us that composition effects do operate, but that they can be masked by offsetting forces when studied over longer horizons. While Hendricks and Schoellman (2014) suggest some possible forces that operated in the U.S., it is not clear that these forces generalize to Korea or other advanced economies. This is an open question for future work.

4.0 INPUT PRICE DISCRIMINATION, INFORMATIVE ADVERTISEMENT EFFORTS, AND WELFARE[1]

4.1 INTRODUCTION

The literature on antitrust legislation and price protection policy has a long history extending from the amendment to the Robinson-Patman Act in the early 1930s, which originally aimed to protect small firms from potential predatory or anti-competitive behaviors by large competitors in intermediate goods markets. But despite the original purpose of the Robinson-Patman Act, as well as the extensive literature on price discrimination, price discrimination with *interdependent* demand did not receive sufficient attention until Katz (1987) recognized its significance in the intermediate goods market. Furthermore, although many papers have studied the interdependency of the input demand since Katz (1987),[2] due to their sole focus on the downstream firms' decisions of strategic substitutes,[3] they have argued that third-degree price discrimination by an upstream monopolist that charges more efficient downstream sellers higher prices hurts social welfare.[4] In contrast, this paper

[1]This research is a joint work with Seung-Gyu Sim.

[2]DeGraba (1990) finds that allowing third-degree price discrimination dampens downstream firms' investment in cost-saving technology, and Yoshida (2000) shows that an increase in aggregate output under third-degree price discrimination is a sufficient condition for welfare deterioration in an intermediate goods market, whereas it is a necessary condition for welfare improvement in a final goods market. Valletti (2003) generalizes the results that Yoshida (2000) obtained.

[3]For the formal definition of strategic substitutes and complements, see Bulow et al. (1985). In our paper, the advertisement effort by one seller increases the rival's value of marginal product, and, hence, production. It causes the complementarity between the advertisement efforts and quantity decisions, although it is different from "strategic complementarity" by Bulow et al. (1985).

[4]Although some recent studies, such as Herweg and Müller (2012), Dertwinkel-Kalt et al. (2015b), Dertwinkel-Kalt et al. (2015a), and Kim and Sim (2015), report the reversed welfare implication of allowing third-degree price discrimination, those studies also rely on the same driving force that a more efficient seller is charged more–because they presume downstream sellers' decisions to be strategic substitutes.

presents a reversal of the welfare implications in the previous literature by incorporating the downstream sellers' advertisement or commercialization efforts into the standard framework by Katz (1987), DeGraba (1990), and Yoshida (2000).

Usually, downstream sellers purchase inputs from a supplier, produce the final products, and advertise them. Unless the final products are sufficiently differentiated, the independent advertisement decision by each seller causes the free-rider problem by its rivals. For example, consider two sellers retailing the same products in different ways. One seller builds up an online store, whereas the other seller manages an offline store at a higher cost than the former. Both sellers put their own advertisement efforts to attract potential customers. The key assumption is that the commodity served by each seller is identical and delivers the same level of utility to buyers. Potential customers observing one seller's advertising may purchase the commodity from the other seller. Without proper compensation for those advertising efforts, both sellers may look for the opportunity to free-ride on the rival's efforts.

This paper analyzes a monopolistic input market in which a monopolistic supplier provides two downstream sellers with inputs, and those sellers simultaneously choose their own advertising efforts and, then, output levels. In a sharp contrast to DeGraba (1990), Yoshida (2000), Dertwinkel-Kalt et al. (2015a), and Kim and Sim (2015), which argue that efficient sellers are charged higher unit prices under third-degree price discrimination, this paper demonstrates that when the advertisement cost is sufficiently inelastic to the advertisement efforts, allowing third-degree price discrimination leads the supplier to offer one seller (cost-efficient seller if they are asymmetric) a price discount[5] and charge the other (inefficient seller) a higher fee, which, in turn, induces the former (the latter) to do more (less) advertising. In fact, the supplier encourages the aggregate advertisement efforts and, hence, production by clarifying the claim on the enhanced demand and penalizing free-riding behavior. Furthermore, the price discount given to the efficient seller in the asymmetric case

[5]Indeed, Intel, for the purpose of resolving the under-advertisement problem by computer manufacturers, practiced a marketing strategy such that it had offered advertisement rebates to the computer manufacturers that put "intel inside" in their own advertisements. The price discount in this paper can be understood as a similar marketing rebate. However, Intel was charged by both the U.S. antitrust authorities and the European Commission (EC) which claimed that Intel's marketing rebate to PC producers was a predatory strategy to maintain its market share against AMD and a price-discrimination strategy to favor relatively larger manufacturers. Since we assume a monopolistic supplier, our aim is not to evaluate Intel's marketing strategy. For more detailed information, see Lee et al. (2013).

improves allocation efficiency, as well. As a result, when the advertisement cost is sufficiently inelastic, allowing third-degree price discrimination can improve social surplus, which is also a reversal of the welfare implications in the previous literature.

The reversed welfare implication in the current paper is consistent with the findings by Inderst and Valletti (2009), Inderst and Shaffer (2009) and Arya and Mittendorf (2010). Inderst and Valletti (2009) show that when input prices are constrained by the threat of demand-side substitution, a more efficient seller is charged a lower price. Also, Inderst and Shaffer (2009) show that when the monopolistic supplier offers two-part tariff contracts, the supplier offers a lower price to more efficient firms. The main emphasis of this paper is to demonstrate that the upstream firm may want to treat otherwise identical downstream firms differently. This is consistent with the previous contract theory papers such as Levitt (1995) and Salant and Shaffer (1999) that even if agents are symmetric, asymmetric contracts can be optimal for the principal.

The paper proceeds as follows. In Section 2, we review the benchmark model and develop a model with advertisement in which sellers are symmetric. Moreover, we set the stage for simultaneous advertisement decisions between the supplier's and sellers' decision nodes. Section 3 extends our scope to include the asymmetric case. Section 4 discusses the case in which the advertisement efforts are controlled by the supplier and Section 5 concludes.

4.2 THE MODEL

4.2.1 The Benchmark Model without Advertisement

Consider a vertically related market structure in which one monopolistic supplier (hereafter 'supplier') provides two downstream final goods sellers ('sellers') with intermediate goods. Each seller, denoted by $i \in \{1, 2\}$, purchases intermediate goods at unit price of $k_i \in \mathbb{R}_+$, assembles one unit of final goods at cost c_i using one unit of intermediate goods, and sells those final goods *à la* Cournot competition. Throughout the paper, (c_1, c_2) are public information. The inverse demand for final goods is given by $P(q_1 + q_2) = a - b(q_1 + q_2)$.

While b is a positive constant in the benchmark model, it will be generalized in the next subsection. Without loss of generality, it is assumed that $c_1 \leq c_2$. The setting is equivalent to the canonical framework by Katz (1987), DeGraba (1990), and Yoshida (2000). When a is not large enough, we may get an equilibrium in which only one seller stays in the market. Since it is not of our interest, we assumes that a is sufficiently large such that both sellers produce in equilibrium.

Following Katz (1987), DeGraba (1990), and Yoshida (2000), we posit that the supplier publicly posts (k_1, k_2) up front.[6] Seller $i \in \{1, 2\}$, taking both input price $k_i \in \mathbb{R}_+$ and rival's choice $q_{-i} \in \mathbb{R}_+$ as given, chooses its best response $\tilde{q}_i(k_i; q_{-i}) \in \mathbb{R}_+$ to maximize $(a - b(\tilde{q}_i(k_i; q_{-i}) + q_{-i}) - k_i - c_i)\tilde{q}_i(k_i; q_{-i})$. The mutual best responses imply that, given $(k_1, k_2) \in \mathbb{R}^2_+$, seller i's optimal decision, $q_i(k_i; k_{-i}) \in \mathbb{R}_+$, should solve for

$$q_i(k_i; k_{-i}) = \tilde{q}_i(k_i; \tilde{q}_{-i}(k_{-i}; q_i(k_i; k_{-i}))). \tag{4.1}$$

Then, the equilibrium markup of seller $i \in \{1, 2\}$ is given by

$$m_i(k_i, k_{-i}) = \frac{1}{3}[a - 2c_i - 2k_i + c_{-i} + k_{-i}], \tag{4.2}$$

and the profit of seller i is given by $m_i(k_i, k_{-i})q_i(k_i, k_{-i})$. When third-degree price discrimination is allowed, the supplier chooses $(k_1, k_2) \in \mathbb{R}^2_+$ such that

$$\begin{aligned} \max_{k_1,\ k_2}\ & k_1 q_1(k_1; k_2) + k_2 q_2(k_2; k_1) \\ s.t.\ \ & q_i(k_i; k_{-i}) = q_i(k_i; q_{-i}(k_{-i}; q_i(k_i; k_{-i}))) \ \text{ for each } \ i \in \{1, 2\}. \end{aligned} \tag{4.3}$$

When price discrimination is prohibited, the supplier solves for the same profit-maximizing problem with one additional constraint, $k_1 = k_2$.

Lemma 1. *When third-degree price discrimination is allowed,*

(i) the supplier charges a higher unit price to the cost-efficient seller, and

(ii) social welfare deteriorates due to the allocation inefficiency.

In particular, if $c_1 = c_2$, the supplier charges the same price to both sellers and implements the same symmetric equilibrium regardless of the pricing regime.

[6] By relaxing this assumption, McAfee and Schwartz (1994) and Rey and Vergé (2004) study how firms' beliefs react based on their rival's contract.

The underlying intuition of Lemma 1, which is a replication of Katz (1987), DeGraba (1990), and Yoshida (2000), is that allowing third-degree price discrimination discourages (encourages) production by the efficient (inefficient) seller because the supplier charges a higher (lower) price to the efficient (inefficient) seller. Unless total output increases significantly, social welfare deteriorates due to the allocation inefficiency.[7] If the sellers are symmetric, all those papers predict that allowing third-degree price discrimination has no impact on the equilibrium outcome.

4.2.2 A Symmetric Model with Advertisement

In this subsection, we develop a simple model with advertisement in which sellers are symmetric, i.e., $c_1 = c_2$. Later, we extend our model to include the asymmetric case in Section 4.3.

The inverse demand curve adopted in the previous subsection is rationalized in the presence of a fixed number of homogeneous consumers having quasi-linear preferences over the homogeneous final goods.[8] Instead, this subsection endogenizes the number of consumers to be determined by aggregate advertisement efforts. Suppose that a 'representative' consumer consuming $\hat{q}$-units of the final goods has the following quasi-linear preferences:

$$u(\hat{q}) = a\hat{q} - \frac{b}{2}\hat{q}^2 - P\hat{q}, \tag{4.4}$$

where P is the market price of the final goods as before; a is positive and sufficiently large; and b is also positive.[9] The individual demand for the final goods is obtained by $\hat{q} = (a-P)/b$. Let e_i be the advertisement effort of seller $i \in \{1, 2\}$ and $x(e_1, e_2)$ the number of consumers who purchase the final goods from either seller[10] after observing the advertisement. Then,

[7] This is the case because the discussion of 'α-efficiency' in Yoshida (2000) is dropped.

[8] This paper focuses on the effect of 'demand enhancement' by advertisement, assuming the homogeneous goods. One can consider another potential effect of 'business stealing,' which makes the advertisement decision to be strategic substitutes. Although the latter effect can be analyzed, together with the former, in a price competition model with product differentiation, we confine our attention to the (qualitative) interaction among the strategic complementary advertisement decisions by the sellers and input price discrimination by the supplier, and leave the 'business stealing issue' for future (quantitative) research.

[9] Alternatively, we can assume quadratic preferences over two goods with the usual budget constraint. The first goods is of our interest and the other goods can be the numéraire.

[10] For the case with market segmentation, see Inderst and Shaffer (2009).

the aggregate demand is given by $Q = x(e_1, e_2)(a - P)/b$, and the aggregate inverse demand function by $P = a - bQ/x(e_1, e_2)$ where it should be the case that $Q = q_1 + q_2$ in equilibrium. For simplicity, let us assume that $x(e_1, e_2) = e_1 + e_2$. The advertisement effort entails the convex cost structure, $\phi(e_i) = e_i^\eta/\eta$, where $\eta > 1$.[11] After observing (k_1, k_2), both sellers make their own advertisement efforts simultaneously and then choose their outputs at the same time.

In stage 3, seller i, taking (k_1, k_2, e_1, e_2) and the rival's $q_{-i} \in \mathbb{R}_+$ as given, chooses $q_i(e_i, e_{-i}, k_i, k_{-i}; q_{-i}) \in \mathbb{R}_+$ such that

$$q_i(e_i, e_{-i}, k_i, k_{-i}; q_{-i}) \in \underset{q_i}{\operatorname{argmax}} \ (a - (b/x(e_1, e_2))(q_i + q_{-i}) - k_i - c_i)q_i - \phi(e_i), \tag{4.5}$$

where $c_1 = c_2 = c$. The mutual best responses jointly define $q_i(e_i, e_{-i}, k_i; k_{-i}) \in \mathbb{R}_+$ such that for each $i \in \{1, 2\}$, $q_i(e_i, e_{-i}, k_i; k_{-i}) = \tilde{q}_i(e_i, e_{-i}, k_i, k_{-i}; \tilde{q}_{-i}(e_{-i}, e_i, k_{-i}, k_i; q_i(e_i, e_{-i}, k_i; k_{-i})))$. In particular, the linear demand structure implies that

$$q_i(e_i, e_{-i}, k_i; k_{-i}) = x(e_1, e_2)q_i(k_i; k_{-i}), \quad \text{where} \quad q_i(k_i; k_{-i}) \quad \text{solves for} \quad (4.1). \tag{4.6}$$

Given (e_1, e_2, k_1, k_2), the sales profit for seller $i \in \{1, 2\}$ is obtained by

$$m_i(k_i, k_{-i})x(e_1, e_2)q_i(k_i; k_{-i}), \tag{4.7}$$

where $m_i(k_i, k_{-i})$ is given by (4.2).

In stage 2, each seller chooses the optimal advertisement effort $e_i(k_i, k_{-i}) \in \mathbb{R}_+$ such that

$$e_i(k_i, k_{-i}) \in \underset{e_i}{\operatorname{argmax}} \ m_i(k_i, k_{-i})x(e_1, e_2)q_i(k_i; k_{-i}) - \phi(e_i), \tag{4.8}$$

which implies that

$$e_1(k_1, k_2) + e_2(k_2, k_1) = \left(\frac{1}{b}\right)^{\frac{1}{\eta-1}} \left(m_1(k_1, k_2)^{\frac{2}{\eta-1}} + m_2(k_2, k_1)^{\frac{2}{\eta-1}}\right). \tag{4.9}$$

Lemma 2 highlights the key role of advertisement in the model. In the benchmark model, the sellers compete à la Cournot, so that an increase in production by one seller reduces his rival's profit. Also, the technology investment by one seller in DeGraba (1990)

[11]The convexity requires that η should be greater than one.

and Choi (1995), being strategic substitutes, reduces the other's market share and, hence, profit. However, seller i's advertisement effort in our model increases his rival's as well as his own sales profit, unless the final goods are differentiated sufficiently. Without proper compensation for those advertising efforts due to the equal access to the final goods market, the opportunistic behavior of each seller results in under-advertisement, as in the tragedy of the commons.[12]

Lemma 2. *When a seller puts more effort into advertising, his and the other seller's sales profit increases.*

Proof. See Appendix C.

Plugging (4.9) into (4.6) yields

$$x(e_1, e_2)q_i(k_i, k_{-i}) = \left(\frac{1}{b}\right)^{\frac{\eta}{\eta-1}} \left(m_1(k_1, k_2)^{\frac{2}{\eta-1}} + m_2(k_2, k_1)^{\frac{2}{\eta-1}}\right) m_i(k_i, k_{-i}), \tag{4.10}$$

where $e_1 = e_1(k_1, k_2)$ and $e_2 = e_2(k_2, k_1)$. Given (k_1, k_2), the monopolistic supplier's profit is obtained by $\sum_i x(e_1, e_2)q_i(k_i; k_{-i})k_i$, where $x(e_1, e_2)q_i(k_i; k_{-i})$ is presented in (4.10). Denote by superscript D and U the outcome under the discriminatory and the uniform pricing rule, respectively. When third-degree price discrimination is allowed, the supplier chooses (k_1^D, k_2^D) such that

$$(k_1^D, k_2^D) \;\; \in \;\; \underset{k_1,\ k_2}{\text{argmax}} \;\; k_1 x(e_1, e_2)q_1(k_1; k_2) + k_2 x(e_1, e_2)q_2(k_2; k_1), \tag{4.11}$$

subject to (4.10). Under the uniform pricing rule, the supplier solves for the same profit-maximizing problem with one additional constraint, $k_1 = k_2 = k^U$.

Next, we show that allowing third-degree price discrimination leads the supplier to charge different input prices for the symmetric sellers. Suppose to the contrary that (k^U, k^U) is an optimal price. Let us define function $f : \mathbb{R}_+^2 \to \mathbb{R}_+$ such that

$$f(k_1, k_2) = k_1 x(e_1, e_2)q_1(k_1; k_2) + k_2 x(e_1, e_2)q_2(k_2; k_1). \tag{4.12}$$

[12] The decentralized economy in which property rights or ownership –especially on the common or public goods –are ill defined results in under-investment on those productive assets compared to the centralized economy. The definitions of common goods and public goods are rivalrous but non-excludable and both non-rivalrous and non-excludable, respectively.

Then, the second order condition requires the corresponding Hessian matrix should be negative definite so that satisfy $\mathbf{k}\mathbf{D}^2 f(\mathbf{k})\mathbf{k}' < 0$, where $\mathbf{k} = (k^U, k^U)$. So, for (k^U, k^U) to be a local maximizer, it must satisfy that $f_{11}(k^U, k^U) < 0$, $f_{22}(k^U, k^U) < 0$, and $f_{11}(k^U, k^U) f_{22}(k^U, k^U) - f_{12}(k^U, k^U) f_{21}(k^U, k^U) > 0$, where f_{ij} denotes the second partial derivative of f with respect to its ith argument, then with respect to it jth argument. It cannot be the case when $\eta \leq 2$. [13] Panels (a), (b) and (c) in Figure 10 show the quantitative results based on different values of $\eta \in \{1.5, 2.0, 2.5\}$. It reveals that when $\eta \leq 2$, the symmetric solution of two first order conditions is not a local maximum, but a saddle point as shown in panels (a) and (b) in Figure 10. It can be a local maximum, in case that $\eta > 2$.

Lemma 3. *When $c_1 = c_2 = c$,*

(i) under the uniform pricing rule, $k^U = (a - c)(\eta - 1)/(2\eta)$, and

(ii) if $\eta \leq 2$, $k_1^D = k_2^D = k^U$. Otherwise, $k_1^D \neq k_2^D$.

Proof. See C.

Lemma 3 tells us that it cannot be optimal to charge the same input prices to both sellers, even when they are symmetric.[14] It is an interesting reversal to the prediction in the previous literature that the monopolistic supplier charges the same prices to those symmetric sellers in the absence of advertisement. Intuitively, the free-riding incentives adversely affect advertisement competition, if both sellers have equal access to the final goods market. In order to avoid the under-advertisement problem, the supplier optimally differentiates the input prices and stimulates one seller's incentive to make an advertising effort. Without loss of generality, we call 'seller 1' the seller who gets the price discount in the symmetric case.

Lemma 4. *Suppose that $c_1 = c_2 = c$. Then, for cost elasticity of advertisement is small enough, i.e., $\eta \leq 2$, the following statements hold:*

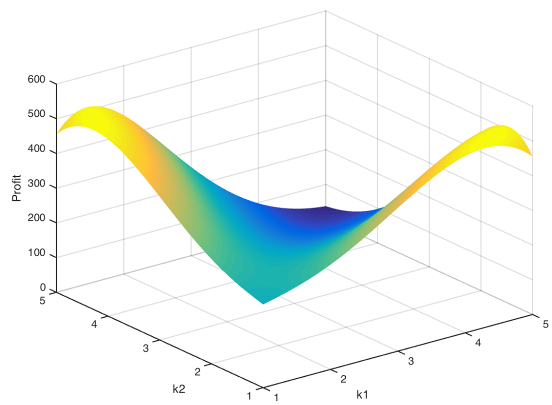
(i) $e_1^D > e_1^U$ and $x(e_1^D, e_2^D) > x(e_1^U, e_2^U)$.

[13] For more detailed proof, see the proof of Lemma 3 in C.

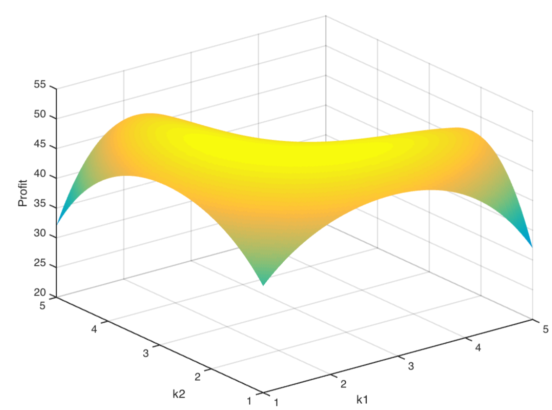
[14] Levitt (1995) shows that in contract theory, asymmetric contracts can be optimal for the principal even if agents are symmetric. In order to analyze this, he set up the model in which the principal's profit depends on the maximum of two agents' outputs. The analytical proof with a maximum function is presented in Levitt (1995).

Figure 10: Profit of monopolistic supplier

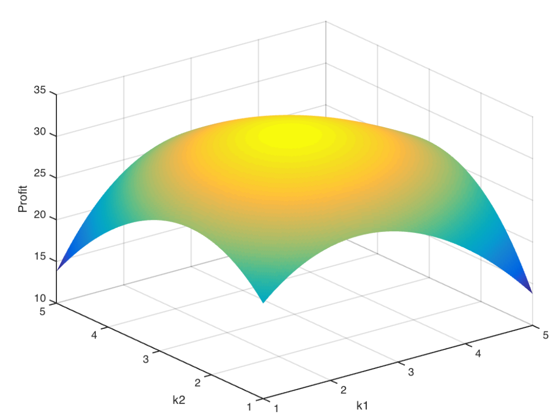
(a) $\eta = 1.5$



(b) $\eta = 2.0$



(c) $\eta = 2.5$



(ii) *Allowing third-degree price discrimination increases (decreases) the market share of seller 1 (seller 2).*

(iii) *If η is sufficiently close to one, $q_1^D > q_1^U$ and $q_1^D + q_2^D > q_1^U + q_2^U$.*

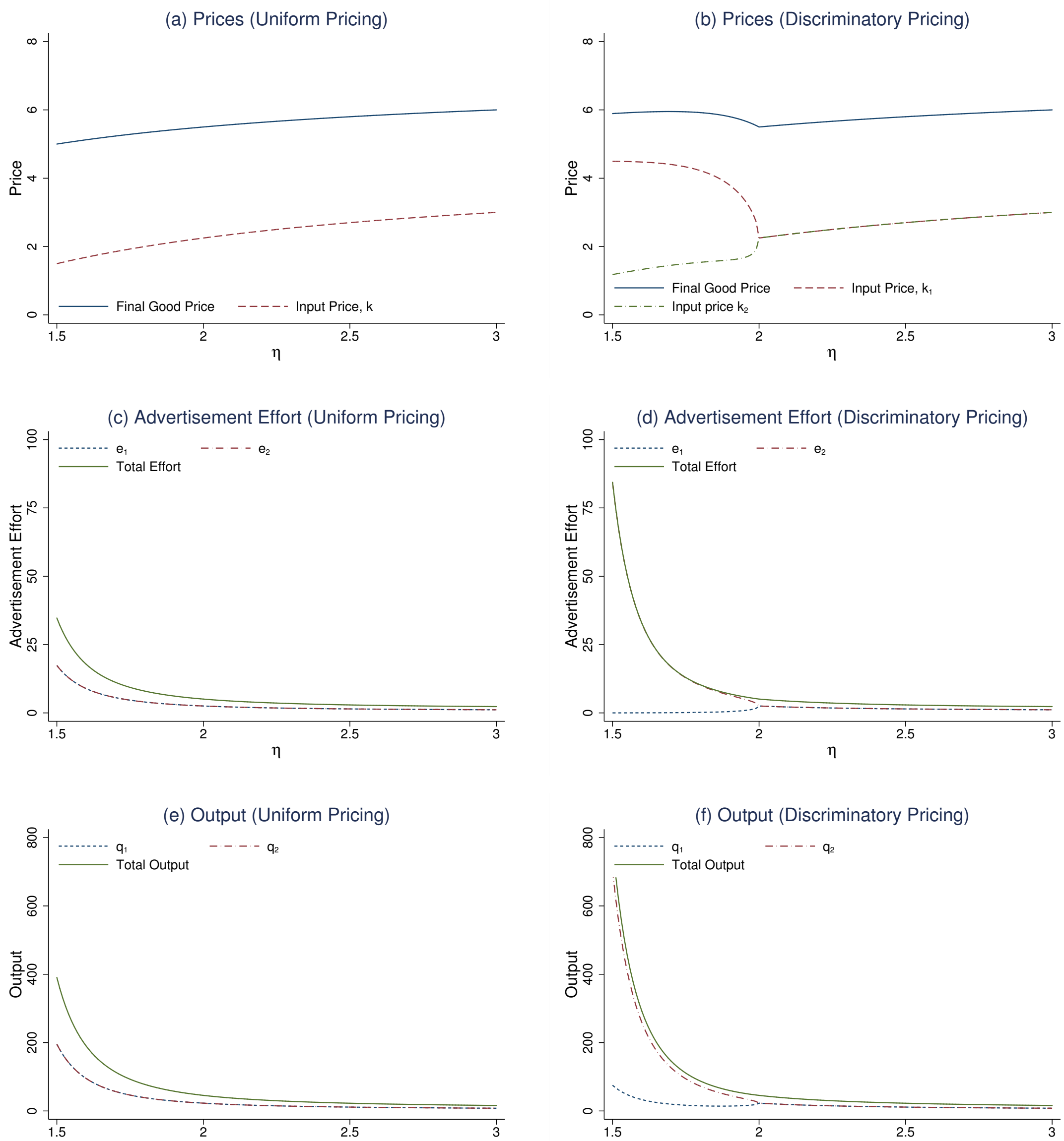
Proof. See C.

When we consider the case in which each agent is symmetric in terms of marginal cost, the input price under the uniform pricing regime can be expressed as a convex combination of the two input prices (k_1^D, k_2^D) under the discriminatory pricing regime. Lemma 4 states that even in the symmetric case, the supplier chooses (k_1^D, k_2^D) as a means of transferring the riding fee from the seller who devotes less effort to the efficient seller who devotes more effort to clarify the latter's claim on the enhanced demand. It also enlarges the aggregate advertising efforts $x(e_1^D, e_2^D)$ by encouraging e_1^D substantially. The price discount given to seller 1 raises the market share of the seller under the discriminatory pricing rule. Apparently, the increases in the aggregate advertisement efforts and the market share of seller 1 induce that $q_1^D > q^U$. But, the output of seller 2 behaves in an ambiguous way because his market share declines. According to our numerical experiments with various parameter values, the aggregate output is more likely to increase when third-degree price discrimination is allowed. In those cases, we obtain the welfare implications such that the profit of seller 1 as well as the supplier increases. However, since the consumer surplus relies on parameter values, we postpone the analysis to the next section.

4.2.3 Numerical Analysis

We show by example that in the presence of advertisement, allowing third-degree price discrimination raises the monopolistic supplier's profit, consumer surplus, and total surplus together, assuming $c_1 = c_2 = 1$. Let $a = 10$ and $b = 1$, which makes $P = 10 - (q_1 + q_2)/(e_1 + e_2)$. Given the demand structure, we analyze the equilibrium strategy and payoff of each player, varying the cost elasticity of advertisement, η, from 1.5 to 3. In particular, we remark that as η increases, all equilibrium outcomes are consistent with the prediction in the previous literature.

Figure 11: Equilibrium Prices and Advertisement Efforts between Symmetric Sellers



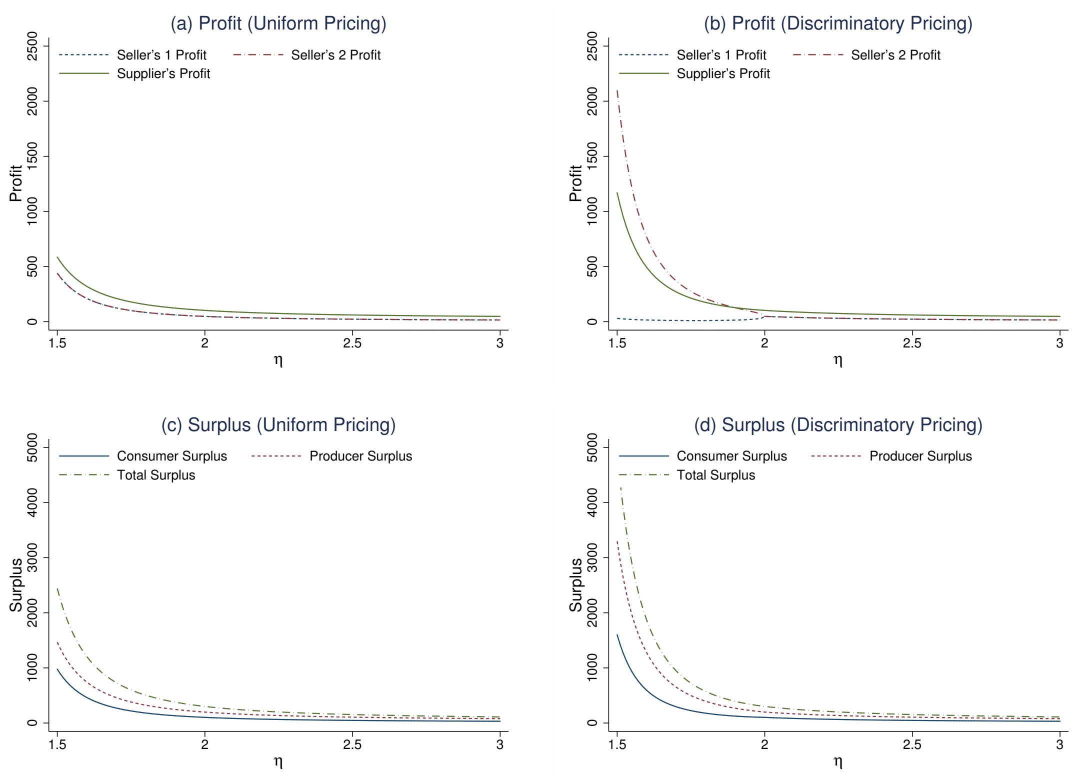
This figure draws the loci of the equilibrium strategy by each player along the cost elasticity of advertisement on the horizontal axis.

Figure 11 presents the equilibrium prices, advertisement efforts and quantities under the uniform and discriminatory pricing regimes. Panels (a), (c), and (e) show each equilibrium outcome under the uniform pricing rule, whereas panels (b), (d), and (f) show each outcome under the discriminatory pricing rule. Comparison between panels (a) and (b) uncovers an interesting point: the input prices, regardless of the pricing rule, run together in the same direction when cost of advertisement is more elastic, i.e., η increases, but the input prices offered under the discriminatory pricing rule run in the opposite direction when η is small enough.[15] Panels (c)–(f) report that, as a result of the discriminatory treatment by the supplier, the seller who gets the discounted offer puts more effort into advertisement and takes a larger market share, whereas his rival puts less efforts and takes a smaller market share. When the advertising cost becomes less elastic, allowing third-degree price discrimination enables the monopolistic supplier to give one seller a dominant market share in the downstream market. By offering different input prices to those symmetric sellers, the supplier can effectively encourage not only the beneficiary's advertisement effort, but also the aggregate advertisement effort, especially when the advertising effort is less elastic. This, in turns, increases the equilibrium output, as well.

Figure 12 summarizes the equilibrium payoffs of the supplier, sellers, and consumers across different level of cost elasticity of advertisement. Panels (a) and (b) provide supporting evidence for our discussion in the previous section. The supplier and seller 1 with the discounted offer, benefit from switching to the discriminatory pricing regime. Also, panels (c) and (d) show that especially when advertising cost is less elastic, consumer surplus and total surplus dramatically increase due to the regime switch. In fact, under both pricing regimes, as the advertising cost becomes less elastic, all economic agents except seller 2 get a higher surplus, on average, but they benefit more under discriminatory pricing. This example tells us that, as in Coase (1960), by clarifying the claim on or accessibility to the demand enhanced by advertisement, the supplier can increase both social surplus and her own profit.

[15] Note that charging the same prices for both sellers is the local maximum for $\eta \geq 2$. See Section 4.2.2.

Figure 12: Equilibrium Payoffs and Surplus



This figure draws the loci of the equilibrium strategy by each player along the cost elasticity of advertisement on the horizontal axis.

4.3 EXTENSION: ASYMMETRIC SELLERS

4.3.1 An Asymmetric Model with Advertisement

This subsection extends our scope to include the asymmetric case with $c_1 < c_2$. To obtain analytical proofs in the asymmetric case, we add additional assumption that $\eta = 2$. Instead, we extend the range of η through the numerical experiments in the next subsection and show that the main arguments in this section still hold.

Assumption 1 $\eta = 2$.

From the first order conditions of the supplier's problem with respect to k_1 and k_2 together with Assumption 1, we derive the equilibrium pricing strategy as

$$-(4m_1^D - 2m_2^D)(k_1^D m_1^D + k_2^D m_2^D) + ((m_1^D)^2 + (m_2^D)^2)(3m_1^D - 2k_1^D + k_2^D) = 0, \tag{4.13}$$

and

$$-(4m_2^D - 2m_1^D)(k_1^D m_1^D + k_2^D m_2^D) + ((m_1^D)^2 + (m_2^D)^2)(3m_2^D - 2k_2^D + k_1^D) = 0. \tag{4.14}$$

Combining (4.13) and (4.14) yields

$$\begin{aligned}&[a - 5c_1 - 5k_1^D + 4c_2 + 4k_2^D][a - 2c_2 - 4k_2^D + c_1 + 2k_1^D] \\ &\quad = [a - 5c_2 - 5k_2^D + 4c_1 + 4k_1^D][a - 2c_1 - 4k_1^D + c_2 + 2k_2^D].\end{aligned} \tag{4.15}$$

Define a mapping $g_i : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$g_i(k_i) := \alpha_i - (3a - 9c_i + 6c_{-i})k_i + 6(k_i)^2, \tag{4.16}$$

where $\alpha_i = a^2 + 14c_i c_{-i} + 2ac_{-i} - 4ac_i - 5c_i^2 - 8c_{-i}^2$. Evaluating (4.16) at k_1^D and k_2^D yields

$$\begin{aligned}g_1(k_1^D) \;&:=\; \alpha_1 - (3a - 9c_1 + 6c_2)k_1^D + 6(k_1^D)^2 \\ &= \alpha_2 - (3a - 9c_2 + 6c_1)k_2^D + 6(k_2^D)^2 := g_2(k_2^D),\end{aligned} \tag{4.17}$$

where the above equality immediately comes from (4.15). Note that if $a > \frac{1}{2}(c_1 + c_2)$, then $\alpha_1 > \alpha_2$, otherwise $\alpha_1 < \alpha_2$. Thus, when $a > \frac{1}{2}(c_1 + c_2)$ and $k_1 = k_2 = 0$, $g_1(0) > g_2(0)$. Also,

$$g_1'(k) = 12k - (3a - 9c_1 + 6c_2) < 12k - (3a - 9c_2 + 6c_1) = g_2'(k). \tag{4.18}$$

Let us denote the value of k at which $g_1(k) = g_2(k)$ by k^*. Note that at $\underline{k}_1 = (a - 3c_1 + 2c_2)/4$, $g_1'(\underline{k}_1) = 0$. The presumption of $k_1^D \geq k_2^D$, together with $g_1(\underline{k}_1) > g_2(0)$, implies that $k_1^D \geq k^*$ and $k_2^D \geq k^*$. For any $y \geq g_1(k^*)$, define the inverse mapping of $g_1(\cdot)$ and $g_2(\cdot)$ such that $g_1^{-1}(y^*) = k_1$ and $g_2^{-1}(y^*) = k_2$. Note that $g_1^{-1}(y^*) - g_2^{-1}(y^*)$ is strictly increasing in $y^*(\geq g_1(k^*))$. Then, we can solve for the pair of $(\overline{k}_1, \overline{k}_2)$ at which $g_1(\overline{k}_1) = g_2(\overline{k}_2)$ and $\overline{k}_1 - \overline{k}_2 = c_2 - c_1$. Combining the two equations results in $(\overline{k}_1, \overline{k}_2) = (a - c_1, a - c_2)$.

Figure 13 illustrates possible candidates of equilibrium input prices (k_1^D, k_2^D) by summarizing the above observations. Panels (a) and (b) in Figure 13 describes the case with asymmetric sellers and symmetric sellers, respectively. In case of asymmetric sellers, there are multiple candidates, depending on the value of $g_1(k_1) = g_2(k_2)$, but Lemma 5 shows that there is a unique pair of equilibrium input prices (k_1^D, k_2^D) such that the supplier charges a lower (higher) price to the efficient (inefficient) seller.

Lemma 5. *Suppose that $c_1 < c_2$. Then, the monopolistic supplier charges a higher price to the inefficient seller than the efficient seller–i.e., $k_1^D < k_2^D$.*

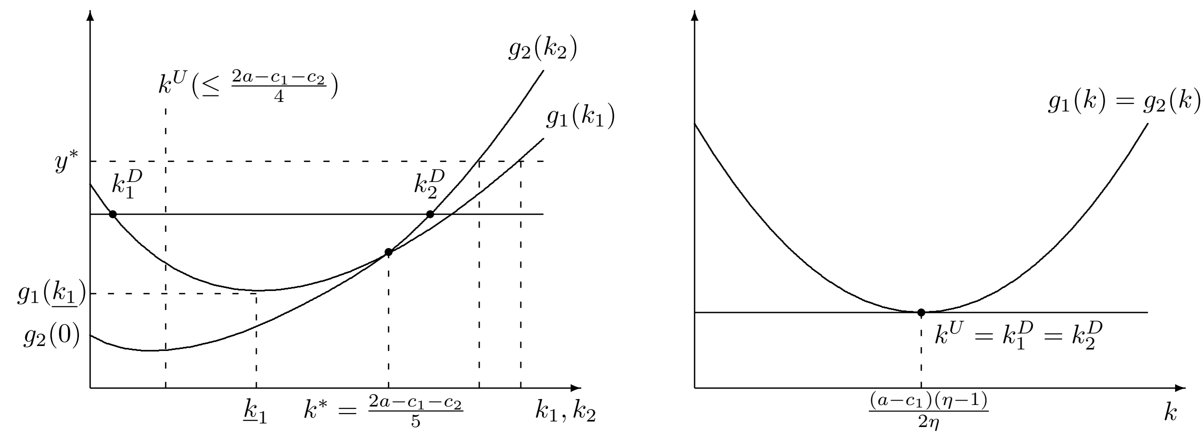
Proof. See C.

Lemma 5 shows that the supplier offers a more favorable price to the efficient seller under the discriminatory pricing rule. The supplier uses price-discrimination to alleviate the free-riding problem associated with the advertisement decision. This result is similar to the main argument of Inderst and Valletti (2009) and Inderst and Shaffer (2009) as well as the one shown in Lemma 4. Intuitively, she chooses (k_1^D, k_2^D) as a means of transferring the riding fee from the inefficient seller to the efficient seller to clarify the ownership of the efficient seller on the enhanced demand.

Lemma 6. *Suppose that $c_1 < c_2$. Then, the following statements hold:*

(i) $e_1^D > e_1^U$ *and* $x(e_1^D, e_2^D) > x(e_1^U, e_2^U)$.

Figure 13: The Optimal Input Prices



(a) Asymmetric Sellers ($c_1 < c_2$) (b) Symmetric Sellers ($c_1 = c_2$)

The horizontal axes of Panel (a) and (b) represent the input prices. In case of symmetric sellers, when $\eta \leq 2$, the solutions are on the saddle point. Hence, the solution in Panel (b) does not hold. See Section 4.2.2.

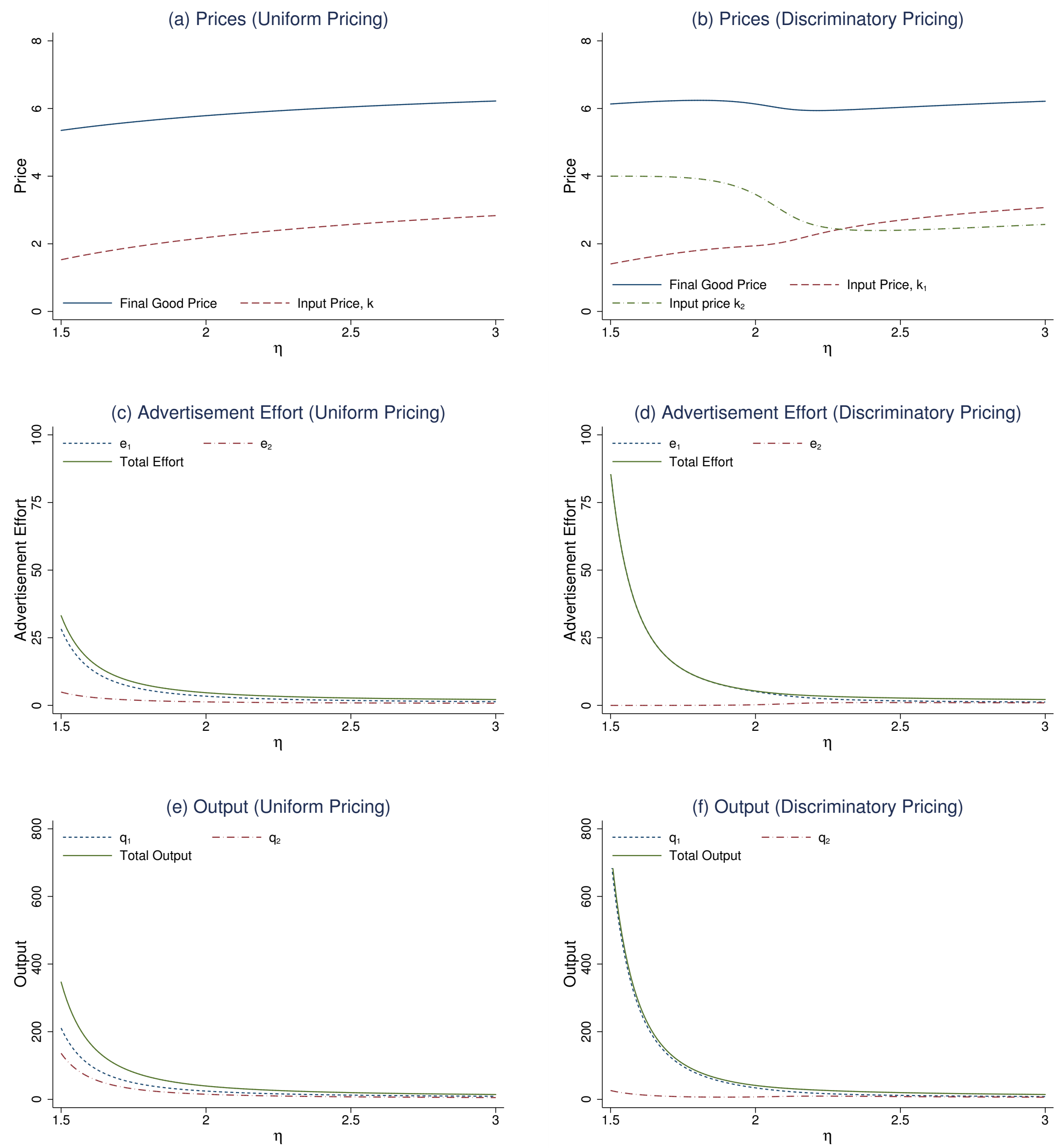
(ii) Allowing third-degree price discrimination increases (decreases) the market share of the cost-efficient seller (inefficient seller).

(iii) $q_1^D > q_1^U$.

Proof. See C.

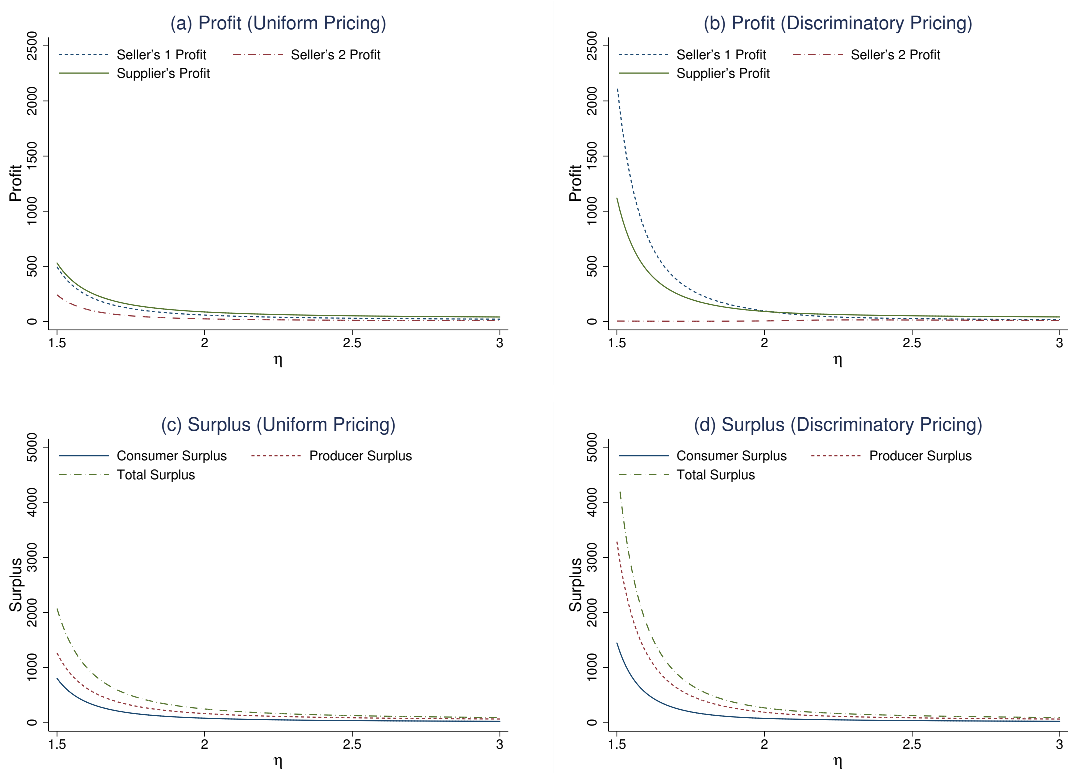
The first statement in Lemma 6 states that the efficient seller puts in more effort and the inefficient seller puts in less effort under the discriminatory pricing rule than under the uniform pricing rule, which is analogous to the first statement in Lemma 4. As a result, we obtain larger aggregate advertisement effort under the discriminatory pricing regime. As shown in Lemma 4, the cost-efficient seller produces more and takes a larger market share when third-degree price discrimination is allowed. But, the output of the inefficient seller is somewhat ambiguous, although it slightly declines in our numerical experiments.

Figure 14: Equilibrium Prices and Advertisement Efforts between Asymmetric Sellers



This figure draws the loci of the equilibrium strategy by each player along the cost elasticity of advertisement on the horizontal axis. At the intersection point in panel (b), ($\eta \approx 2.3$) $k_1^D = k_2^D = k^U$. In panel (d), the advertisement effort e_1 almost overlaps the total effort.

Figure 15: Equilibrium Payoffs and Surplus



This figure draws the loci of the equilibrium strategy by each player along the cost elasticity of advertisement on the horizontal axis.

4.3.2 Numerical Analysis

Figure 14 presents the loci of each equilibrium strategy along the cost elasticity of advertisement η when $c_1 = 1$ and $c_2 = 2$. In panel (a), when $\eta \approx 2.3$, $k_1^D = k_2^D$. By construction, $k_1^D = k_2^D = k^U$ at the intersection point. When the advertising cost is less elastic than the intersection point, $k_1 < k_2$ ($k_1 > k_2$). Moreover, as η decreases, the supplier charges less overall, but she may charge the inefficient seller relatively more under the discriminatory pricing rule. Panels (c) to (f) present the advertisement efforts and outputs under two different pricing regimes. All of these panels reveal patterns similar to those in the symmetric case.

Figure 15 analyzes each player's equilibrium payoffs across different level of cost elasticity of advertisement. The results are the same with the symmetric case.

4.4 DISCUSSION: ADVERTISEMENT REBATE BY THE SUPPLIER

This section considers the case in which the supplier determines the effort level of advertisement by each seller and provides each with an advertisement rebate. In order to maximize her own profit, the supplier chooses $(\tilde{e_1}, \tilde{e_2}, \tilde{k_1}, \tilde{k_2})$ such that

$$\begin{aligned}(\tilde{e_1}, \tilde{e_2}, \tilde{k_1}, \tilde{k_2}) \quad \in \quad & \underset{e_1,e_2,k_1,k_2}{\operatorname{argmax}}\, x(e_1,e_2)[k_1 q_1(k_1;k_2) + k_2 q_2(k_2;k_1)] - \phi(e_1) - \phi(e_2) \\ & s.t. \;\; q_i(k_i;k_{-i}) = q_i(k_i; q_{-i}(k_{-i}; q_i(k_i;k_{-i}))) \;\; \text{for each} \;\; i \in \{1,2\}.\end{aligned} \tag{4.19}$$

Lemma 7. *When the monopolistic supplier can choose each seller's advertisement efforts, the supplier implements a higher level of aggregate advertisement efforts under the discriminatory pricing regime than under the uniform pricing regime. Accordingly, the aggregate output level is higher under the former. Moreover, the supplier charges* $\tilde{k}_2 < \tilde{k}^U < \tilde{k}_1$.

Proof. See C.

Lemma 7 states that by choosing the optimal level of advertisement, the supplier can force the sellers to put more effort into advertisement under discriminatory pricing than they would under the uniform pricing regime. Furthermore, when she can coordinate the seller's advertisement efforts, the supplier charges a higher price to the cost-efficient seller. Interestingly, since the free-rider problem no longer matters, the supplier does not exploit $(\tilde{k}_1, \tilde{k}_2)$ as a means of charging a fee for potential free-riding. Instead, she charges a higher price to the cost-efficient but inelastic seller, which shows that the optimal input prices in both pricing regimes follow the same pattern in Katz (1987), DeGraba (1990), and Yoshida (2000). However, unlike the previous literature, the aggregate output under discriminatory pricing is higher than that under uniform pricing because of the effect of an increase in advertisement effort.

In this case, the welfare implication is somewhat ambiguous. This is because, on the one hand, allowing for third-degree price discrimination induces more advertisement efforts by the supplier, and, on the other hand, more production by the inefficient seller, as in the previous literature. Hence, total surplus may or may not be improved by permitting third-degree price discrimination.

4.5 CONCLUSION

Unless the final products are sufficiently differentiated, the independent advertisement decisions of final goods sellers may result in under-advertisement. Motivated by the free-rider problem in the advertisement stage, this paper examines a vertically related market in which a monopolistic supplier provides two downstream sellers with inputs, and those sellers simultaneously choose their own advertising or commercialization efforts, followed by their own output levels. We find that when the demand for final goods is sufficiently elastic with respect to sellers' advertisement efforts, the monopolistic supplier charges a lower (higher) price to the efficient (inefficient) seller in order to encourage advertisement under the discriminatory pricing rule. As a result, total output, consumer surplus, and social surplus may increase when third-degree price discrimination is permitted. This is a reversal of the welfare implications of banning price discrimination suggested in previous studies, such as those of Katz (1987), DeGraba (1990), and Yoshida (2000), which presume the sellers' decisions to be strategic substitutes.

One potential shortcoming of the present research is that– our focus on the role of complementarity between advertisement and production decision leads us to think of the supplier as a coordinator to mitigate the free-rider problem– and, thus, to neglect potential product differentiation by downstream sellers. When the final goods are differentiated, one seller's advertisement effort may enable him to promote the total number of customers in the market and/or to attract his rival's customers at the same time. We can analyze both effects of 'demand enhancement' and 'business stealing' by developing an alternative model with product differentiation. We leave it to future quantitative research to determine whether this shortcoming has led us to oversimplify our argument and to overvalue the positive impact of third-degree price discrimination.

APPENDIX A

CHAPTER 2

A.1 DATA

To measure the annual labor market transition rates, I define EE transition as job changes in which employed workers leave for a job with a different employer and UE as the proportion of workers who were unemployed in the previous year and are currently employed.

To classify job change, I use the empirical strategy in Stewart (2002). I restrict attention to individuals who are employed during the time of the survey and satisfy at least one of the following conditions: i) the individual had more than one employer in the previous year; ii) the individual experienced more than one spell of unemployment (but, not more than two or fewer weeks of unemployment); and iii) the individual experienced a change in the one-digit industry code.[1]

[1] For more-detailed information on EE transition, see Stewart (2002).

APPENDIX B

CHAPTER 3

B.1 SAMPLE FOR ANALYSIS 2: KAYMAK 2009

To estimate the effect of years of schooling on real wage following Kaymak (2009), we use the yearly data from Korean Survey Report on Wage Structure between 1980 to 2011. The sample is restricted to men between 25- and 60-years-old. Then the sample include cohorts, who were born in the same year, born between 1920-1985.

Wages are measured by hour earnings that are calculated by dividing monthly wage and salary earnings by monthly worked hours. Using CPI in 2010, we convert the wages to real wages. The Korean Survey Report on Wage Structure provides 5 groups of education attainments; Elementary, Middle and High schools, some college and four-years university. We assign 6, 9, 12, 14 and 16 years to each group for years of schooling respectively. We drop workers with less than half of the minimum wages in 2010, $4.11 and we also drop workers who worked less than 35 hours per week.

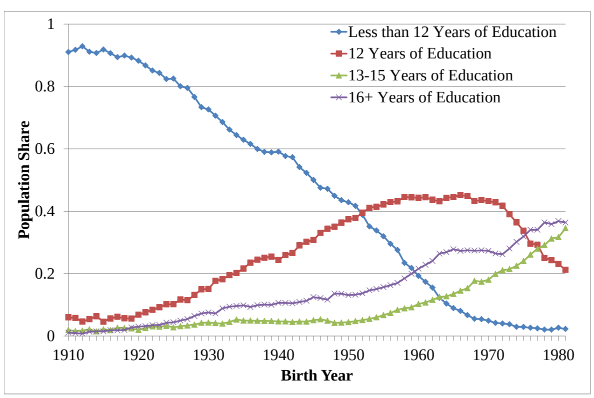
B.2 SAMPLE FOR ANALYSIS 3: JUHN, KIM, AND VELLA 2005

To estimate relation between college wages premium and college share following Juhn et al. (2005), we use the data from both Korean Survey Report on Wage Structure and Korea Population Census. For information about the real wages, we use wage profile from the

Korean Survey Report on Wage Structure from 1980 to 2010 every 5-year. Wages are measured by hour earnings that are calculated by dividing monthly wage and salary earnings by monthly worked hours. Using CPI in 2010, we convert the wages to real wages. We drop workers with less than half of the minimum wages in 2010, \$4.11. and we also drop workers who worked less than 35 hours per week. We get the numbers of college and non-college graduates from the data from Korean population census. We also restrict the sample to men between 27 and 65 years old and women between 23 and 65 years to consider people who under a schooling. To get the numbers of college graduates and non-college graduates for each cohort, we predict college share of each cohort at age 35 as a specified equation provided by Juhn et al. (2005) using Korean population census data. We predicted every 5 years using the data from Korean population census 1966 and from 1970 to 2010.

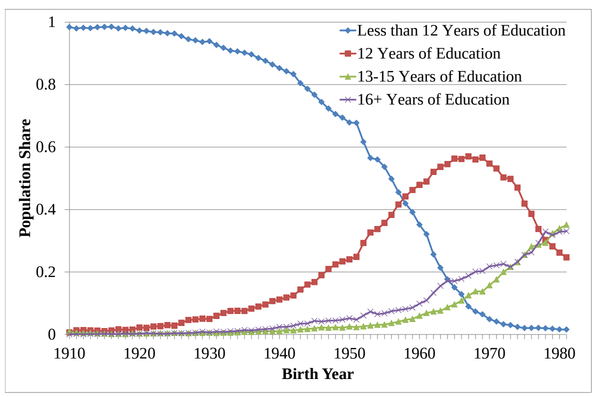
B.3 APPENDIX FIGURE

Figure 16: Educational Attainment by Birth Cohort (Men)



Note: Educational attainment measured in four mutually exclusive and exhaustive categories. Attainment measured using 33-37 year-olds in the 1966 or 1970–2010 Korean population censuses, which provides a unique observation for each birth cohort.

Figure 17: Educational Attainment by Birth Cohort (Women)



Note: Educational attainment measured in four mutually exclusive and exhaustive categories. Attainment measured using 33-37 year-olds in the 1966 or 1970–2010 Korean population censuses, which provides a unique observation for each birth cohort.

APPENDIX C

CHAPTER 4

C.1 PROOFS

Proof of Lemma 2 The seller i's profit is given by

$$\pi_i = \left[a - \frac{b}{(e_1+e_2)}(q_1(e_1,e_2,k_1;k_2) + q_2(e_2,e_1,k_2;k_1)) - k_i - c_i\right] \times q_i(e_i,e_{-i},k_i;k_{-i}) - \phi(e_i), \tag{C.1}$$

where $q_i(e_i,e_{-i},k_i;k_{-i}) = (e_1+e_2)m_i(k_i,k_{-i})/b$. Taking derivative of (C.1) with respect to his rival's advertisement effort yields

$$\frac{\partial \pi_i}{\partial e_{-i}} = [a - (m_1(k_1,k_2) + m_2(k_2,k_1)) - k_i - c_i]m_i(k_i,k_{-i}) > 0. \tag{C.2}$$

Therefore, the advertisement efforts by one seller also increase the other's profit. □

Proof of Lemma 3 Assume that $c_1 = c_2 = c$.

(i) Under the uniform pricing regime, the monopolistic supplier chooses k^U such that

$$2m(k^U,k^U)^{\frac{2}{\eta-1}}\left(m(k^U,k^U) - \frac{\eta+1}{3(\eta-1)}k^U\right) = 0, \tag{C.3}$$

where $m(k^U,k^U) = (a-c-k^U)/3$.

This immediately implies the following optimal input price

$$k^U = \frac{(a-c)(\eta-1)}{2\eta}. \tag{C.4}$$

(ii) The first order optimality condition with respect to k_1^D evaluated at k^U is

$$f_1(k^U, k^U) \quad := \quad 2\left(\frac{a-c-k^U}{3}\right)^{\frac{2}{\eta-1}}\left[\frac{a-c-2k^U}{3} - \frac{2k^U}{3(\eta-1)}\right] = 0. \qquad \text{(C.5)}$$

The value inside the parenthesis is equal to zero, once we plug in the value for k^U, so we get $f_1(k^U, k^U) = 0$. In order to show k^U is an optimal price, we should also check whether $f(\cdot)$ is negative semidefinite at the optimal point. In other words, the second order condition requires Hessian matrix must satisfy $\mathbf{k}\mathbf{D}^2 f(\mathbf{k})\mathbf{k}' \leq 0$, where $\mathbf{k} = (k^U, k^U)$.

The second order condition with respect to k_1^D evaluated at k^U is

$$f_{11}(k^U, k^U) \quad = \quad \frac{1}{\eta^2-1}\left[4^{\frac{1}{1-\eta}+1} \times 3^{\frac{1+\eta}{1-\eta}}\left(\frac{(a-c)(1+\eta)}{\eta}\right)^{\frac{2}{\eta-1}}\right](15-5\eta-2\eta^2). \quad \text{(C.6)}$$

The two roots are $\eta = 0.25(-5-\sqrt{145})$ and $\eta = 0.25(-5+\sqrt{145})$. So, if $\eta > 1.76$, then $f_{11}(k^U, k^U) < 0$. This result shows that second order conditions evaluated at k^U is not always negative semidefinite.

In order to insure that (k^U, k^U) is a symmetric solution, the following condition must be satisfied:

$$\begin{aligned}&f_{11}(k^U, k^U)f_{22}(k^U, k^U) - f_{12}(k^U, k^U)f_{21}(k^U, k^U)\\ &= \frac{\eta}{(\eta-1)^2(\eta+1)}3^{\frac{3+\eta}{1-\eta}}16^{\frac{2-\eta}{1-\eta}}\left(\frac{(a-c)(\eta+1)}{\eta}\right)^{\frac{4}{\eta-1}}(\eta^2+3\eta-10) > 0. \qquad \text{(C.7)}\end{aligned}$$

Hence, η must be strictly greater than 2. $\square$

Proof of Lemma 4 Assume that $c_1 = c_2 = c$.

(i) Note that $m(k,k) = \frac{1}{3}(a-c-k)$. Since $m(-\infty, -\infty) = \infty$ and $m(a-c, a-c) = 0$, $m(k,k)$ monotonically declines with $k \in (-\infty, a-c]$. By construction, $(m(k,k))^{\frac{2}{\eta-1}}$ is strictly decreasing in $k \in (-\infty, a-c]$. Define the threshold $\overline{k} \in \mathbb{R}$ such that

$$(m(k_1^D, k_2^D))^{\frac{2}{\eta-1}} + (m(k_2^D, k_1^D))^{\frac{2}{\eta-1}} = (m(\overline{k}, \overline{k}))^{\frac{2}{\eta-1}} + (m(\overline{k}, \overline{k}))^{\frac{2}{\eta-1}}. \qquad \text{(C.8)}$$

Due to the convex property, we know that $2\overline{k} < k_1^D + k_2^D$.

Conjecture 1 $k^U > \overline{k}$.

Suppose to the contrary that $0 < k^U \leq \overline{k}$. From the supplier's optimal decision and (C.8), we infer that

$$\pi_s(\overline{k}, \overline{k}) \leq \pi_s(k_1^D, k_2^D) \Longrightarrow (m(\overline{k}, \overline{k}) + m(\overline{k}, \overline{k}))\overline{k} \leq m_1^D k_1^D + m_2^D k_2^D. \tag{C.9}$$

Otherwise, $\pi_s(\overline{k}, \overline{k}) > \pi_s(k_1^D, k_2^D)$. Since $k^U \leq \overline{k}$, the first order condition (in stage 3) evaluated at $k = \overline{k}$ implies that

$$\begin{aligned} -\frac{2}{\eta - 1}(m(\overline{k}, \overline{k})^{\frac{2}{\eta-1}-1} + m(\overline{k}, \overline{k})^{\frac{2}{\eta-1}-1})(m(\overline{k}, \overline{k}) + m(\overline{k}, \overline{k}))\overline{k} \\ + ((m(\overline{k}, \overline{k}))^{\frac{2}{\eta-1}} + (m(\overline{k}, \overline{k}))^{\frac{2}{\eta-1}})(3m(\overline{k}, \overline{k}) + 3m(\overline{k}, \overline{k}) - 2\overline{k}) \leq 0. \end{aligned} \tag{C.10}$$

Rewriting (C.10) yields

$$\begin{aligned} \frac{2(m(\overline{k}, \overline{k})^{\frac{2}{\eta-1}-1} + m(\overline{k}, \overline{k})^{\frac{2}{\eta-1}-1})}{(\eta - 1)(3m(\overline{k}, \overline{k}) + 3m(\overline{k}, \overline{k}) - 2\overline{k})} &\geq \frac{(m(\overline{k}, \overline{k}))^{\frac{2}{\eta-1}} + (m(\overline{k}, \overline{k}))^{\frac{2}{\eta-1}}}{\overline{k}m(\overline{k}, \overline{k}) + \overline{k}m(\overline{k}, \overline{k})} \\ \geq \frac{(m_1^D)^{\frac{2}{\eta-1}} + (m_2^D)^{\frac{2}{\eta-1}}}{m_1^D k_1^D + m_2^D k_2^D} &= \frac{2((m_1^D)^{\frac{2}{\eta-1}-1} + (m_2^D)^{\frac{2}{\eta-1}-1})}{(\eta - 1)(3m_1^D + 3m_2^D - k_1^D - k_2^D)}. \end{aligned} \tag{C.11}$$

The first inequality immediately follows from (C.10) and the second inequality from (C.8) and (C.9). The last equality of (C.11) comes from the first order conditions of the supplier's profit maximization problem when third-degree price discrimination is allowed. However, since $2\overline{k} < k_1^D + k_2^D$ by construction,

$$\begin{aligned} \frac{2(m(\overline{k}, \overline{k})^{\frac{2}{\eta-1}-1} + m(\overline{k}, \overline{k})^{\frac{2}{\eta-1}-1})}{(\eta - 1)(3m(\overline{k}, \overline{k}) + 3m(\overline{k}, \overline{k}) - 2\overline{k})} &= \frac{4((2a - 2c - 2\overline{k})/3)^{\frac{2}{\eta-1}-1}}{(\eta - 1)(2a - 2c - 4\overline{k})} \\ < \frac{4((2a - 2c - (k_1^D + k_2^D))/3)^{\frac{2}{\eta-1}-1}}{(\eta - 1)(2a - 2c - 2(k_1^D + k_2^D))} &< \frac{2((m_1^D)^{\frac{2}{\eta-1}-1} + (m_2^D)^{\frac{2}{\eta-1}-1})}{(\eta - 1)(3m_1^D + 3m_2^D - k_1^D - k_2^D)}, \end{aligned} \tag{C.12}$$

which is a contradiction. The second inequality comes from condition $2\overline{k} < k_1^D + k_2^D$, and the last inequality comes from convexity. Therefore, we conclude that $k^U > \max\{\overline{k}, 0\}$.

From the above argument, we know that $m(k^U, k^U) < m(\bar{k}, \bar{k})$. Then,

$$\begin{aligned}
e^U + e^U &= (\eta b)^{\frac{-1}{\eta-1}}[(m(k^U, k^U))^{\frac{2}{\eta-1}} + (m(k^U, k^U))^{\frac{2}{\eta-1}}] \\
&< (\eta b)^{\frac{-1}{\eta-1}}[(m(\bar{k}, \bar{k}))^{\frac{2}{\eta-1}} + (m(\bar{k}, \bar{k}))^{\frac{2}{\eta-1}}] \\
&= (\eta b)^{\frac{-1}{\eta-1}}[(m(k_1^D, k_2^D))^{\frac{2}{\eta-1}} + (m(k_2^D, k_1^D))^{\frac{2}{\eta-1}}] = e_1^D + e_2^D
\end{aligned} \tag{C.13}$$

Moreover, since $m(k^U, k^U) < m(\bar{k}, \bar{k}) < m(k_1^D, k_2^D)$, it is obtained that $e^U < e_1^D$.

(*ii*) The market share of seller 1 is given by

$$\frac{q_1^D}{q_1^D + q_2^D} = \frac{m(k_1^D, k_2^D)}{m(k_1^D, k_2^D) + m(k_2^D, k_1^D)} \quad \text{and} \quad \frac{q^U}{q^U + q^U} = \frac{m(k^U, k^U)}{m(k^U, k^U) + m(k^U, k^U)}. \tag{C.14}$$

Since $m(k_1^D, k_2^D) > m(k^U, k^U)$ and $m(k_2^D, k_1^D) < m(k^U, k^U)$,

$$\begin{aligned}
\frac{m(k_1^D, k_2^D)}{m(k_1^D, k_2^D) + m(k_2^D, k_1^D)} &> \frac{m(k^U, k^U)}{m(k^U, k^U) + m(k_2^D, k_1^D)} \\
&> \frac{m(k^U, k^U)}{m(k^U, k^U) + m(k^U, k^U)}.
\end{aligned} \tag{C.15}$$

Therefore, allowing third-degree price discrimination increases the market share of seller 1 and decreases that of seller 2.

(*iii*) Given (4.6), the ratio of seller 1's quantity under the discriminatory pricing rule to it under the uniform pricing rule is given by

$$\frac{q_1^D}{q_1^U} = \frac{(1 + (m(k_2^D, k_1^D)/m(k_1^D, k_2^D))^{\frac{2}{\eta-1}})m(k_1^D, k_2^D)}{((m(k^U, k^U)/m(k_1^D, k_2^D))^{\frac{2}{\eta-1}} + (m(k^U, k^U)/m(k_1^D, k_2^D))^{\frac{2}{\eta-1}})m(k^U, k^U)}. \tag{C.16}$$

Since $m(k_1^D, k_2^D) > m(k_2^D, k_1^D)$ and $m(k_1^D, k_2^D) > m(k^U, k^U)$,

$$\lim_{\eta \to 1} \left(\frac{m(k_2^D, k_1^D)}{m(k_1^D, k_2^D)} \right)^{\frac{2}{\eta-1}} = 0 \quad \text{and} \quad \lim_{\eta \to 1} \left(\frac{m(k^U, k^U)}{m(k_1^D, k_2^D)} \right)^{\frac{2}{\eta-1}} = 0. \tag{C.17}$$

Then, as η goes to one, the quantity ratio (C.42) will be

$$\lim_{\eta \to 1} \frac{q_1^D}{q_1^U} > 0. \tag{C.18}$$

This implies that when η is sufficiently close to one, $q_1^D > q_1^U$.

Similarly, the ratio of total quantities between two different regimes is given by

$$\frac{q_1^D + q_2^D}{q_1^U + q_2^U} = \frac{(1 + (m(k_2^D, k_1^D)/m(k_1^D, k_2^D))^{\frac{2}{\eta-1}})(m(k_1^D, k_2^D) + m(k_2^D, k_1^D))}{2(m(k^U, k^U)/m(k_1^D, k_2^D))^{\frac{2}{\eta-1}})^{\frac{2}{\eta-1}} 2m(k^U, k^U)}. \tag{C.19}$$

So, as η goes to one, the total quantities ratio (C.42) will be

$$\lim_{\eta \to 1} \frac{q_1^D + q_2^D}{q_1^U + q_2^U} > 0. \tag{C.20}$$

This implies that when η is sufficiently close to one, $(q_1^D + q_2^D) > (q_1^U + q_2^U)$. □

Proof of Lemma 5 Suppose to the contrary that $k_1^D \geq k_2^D$. The first-order conditions of the supplier's constrained profit maximization problem imply that

$$-(4m_1^D - 2m_2^D)(k_1^D m_1^D + k_2^D m_2^D) + ((m_1^D)^2 + (m_2^D)^2)(3m_1^D - 2k_1^D + k_2^D) = 0, \tag{C.21}$$

and

$$-(4m_2^D - 2m_1^D)(k_1^D m_1^D + k_2^D m_2^D) + ((m_1^D)^2 + (m_2^D)^2)(3m_2^D - 2k_2^D + k_1^D) = 0. \tag{C.22}$$

Combining (4.13) and (4.14) yields

$$\begin{aligned}&[a - 5c_1 - 5k_1^D + 4c_2 + 4k_2^D][a - 2c_2 - 4k_2^D + c_1 + 2k_1^D] \\ &\quad = [a - 5c_2 - 5k_2^D + 4c_1 + 4k_1^D][a - 2c_1 - 4k_1^D + c_2 + 2k_2^D]. \end{aligned}\tag{C.23}$$

Let us define the following quadratic function

$$g_i(k_i) := \alpha_i - (3a - 9c_i + 6c_{-i})k_i + 6(k_i)^2, \tag{C.24}$$

where $\alpha_i = a^2 + 14c_i c_{-i} + 2ac_{-i} - 4ac_i - 5c_i^2 - 8c_{-i}^2$. Evaluating (C.24) at k_1^D and k_2^D yields

$$\begin{aligned} g_1(k_1^D) \;&:=\; \alpha_1 - (3a - 9c_1 + 6c_2)k_1^D + 6(k_1^D)^2 \\ &= \alpha_2 - (3a - 9c_2 + 6c_1)k_2^D + 6(k_2^D)^2 := g_2(k_2^D), \end{aligned}\tag{C.25}$$

where the above equality immediately comes from (C.23). Note that if $a > \frac{1}{2}(c_1 + c_2)$, then $\alpha_1 > \alpha_2$, otherwise $\alpha_1 < \alpha_2$. Thus, when $a > \frac{1}{2}(c_1 + c_2)$ and $k_1 = k_2 = 0$, $g_1(0) > g_2(0)$. Also,

$$g_1'(k) = 12k - (3a - 9c_1 + 6c_2) < 12k - (3a - 9c_2 + 6c_1) = g_2'(k). \tag{C.26}$$

Let us denote the value of k at which $g_1(k) = g_2(k)$ by k^*. Note that at $\underline{k}_1 = (a-3c_1+2c_2)/4$, $g_1'(\underline{k}_1) = 0$. The presumption of $k_1^D \geq k_2^D$, together with $g_1(\underline{k}_1) > g_2(0)$, implies that $k_1^D \geq k^*$ and $k_2^D \geq k^*$. For any $y \geq g_1(k^*)$, define the inverse mapping of $g_1(\cdot)$ and $g_2(\cdot)$ such that $g_1^{-1}(y^*) = k_1$ and $g_2^{-1}(y^*) = k_2$. Note that $g_1^{-1}(y^*) - g_2^{-1}(y^*)$ is strictly increasing in $y^*(\geq g_1(k^*))$. Then, we can solve for the pair of $(\overline{k}_1, \overline{k}_2)$ at which $g_1(\overline{k}_1) = g_2(\overline{k}_2)$ and $\overline{k}_1 - \overline{k}_2 = c_2 - c_1$. Combining the two equations results in $(\overline{k}_1, \overline{k}_2) = (a-c_1, a-c_2)$.

Combining (C.21) and (C.22) yields

$$\begin{aligned}
&(2m_1^D - m_2^D)(3m_2^D - 2k_2^D + k_1^D) = (2m_2^D - m_1^D)(3m_1^D - 2k_1^D + k_2^D) \\
&\implies 6m_1^D m_2^D - 3(m_2^D)^2 - 4m_1^D k_2^D + 2m_2^D k_2^D + 2m_1^D k_1^D - m_2^D k_1^D \\
&\qquad = 6m_2^D m_1^D - 3(m_1^D)^2 - 4m_2^D k_1^D + 2m_1^D k_1^D + 2m_2^D k_2^D - m_1^D k_2^D \\
&\implies -3(m_2^D)^2 - 4m_1^D k_2^D - m_2^D k_1^D = -3(m_1^D)^2 - 4m_2^D k_1^D - m_1^D k_2^D \\
&\implies 3(m_2^D)^2 - 3m_2^D k_1^D = 3(m_1^D)^2 - 3m_1^D k_2^D \\
&\implies m_2^D(m_2^D - k_1^D) = m_1^D(m_1^D - k_2^D) \qquad \text{(C.27)}
\end{aligned}$$

Case 1: Suppose $(m_1^D - k_2^D) > 0$ and $(m_2^D - k_1^D) > 0$. Then, $m_1^D < m_2^D$, otherwise $m_2^D(m_2^D - k_1^D) < m_1^D(m_1^D - k_2^D)$. This implies $m_1^D - m_2^D = (c_2 - c_1) - (k_1^D - k_2^D) < 0$. Since $k_1 - k_2 = c_2 - c_1$ at y^* by construction, $k_1^D - k_2^D > c_2 - c_1$ for any $y > y^*$. However, in this case,

$$\begin{aligned}
(m_2^D - k_1^D) &= \frac{1}{3}(a - 2c_2 + c_1 - 2(k_2^D + k_1^D)) \\
&< \frac{1}{3}(a - 2c_2 + c_1 - 2(2a - c_1 - c2)) < 0. \qquad \text{(C.28)}
\end{aligned}$$

This is a contradiction.

Case 2: Suppose that $(m_1^D - k_2^D) < 0$ and $(m_2^D - k_1^D) < 0$. Since

$$\begin{aligned}
|m_2^D - k_1^D| &= |\frac{1}{3}(a - 2c_2 + c_1 - 2k_1^D - 2k_2^D)| \\
&> |\frac{1}{3}(a - 2c_1 + c_2 - 2k_1^D - 2k_2^D)| = |m_1^D - k_2^D|, \qquad \text{(C.29)}
\end{aligned}$$

$m_1^D > m_2^D$, otherwise the equality in (C.27) cannot be satisfied. This implies $m_1^D - m_2^D = (c_2 - c_1) - (k_1^D - k_2^D) > 0$. Hence, $(c_2 - c_1) > (k_1^D - k_2^D)$. In this case,

$$m_2^D(m_2^D - k_1^D) = m_1^D(m_1^D - k_2^D) \tag{C.30}$$

$$\frac{m_2^D}{m_1^D} = \frac{(m_1^D - k_2^D)}{(m_2^D - k_1^D)} \tag{C.31}$$

$$\frac{a - 2c_2 + c_1 - 2k_2^D + k_1^D}{a - 2c_1 + c_2 - 2k_1^D + k_2^D} = \frac{a - 2c_1 + c_2 - 2(k_1^D + k_2^D)}{a - 2c_2 + c_1 - 2(k_1^D + k_2^D)}. \tag{C.32}$$

Note that the left hand side of (C.32) is

$$\frac{a - 2c_2 + c_1 - 2k_2^D + k_1^D}{a - 2c_1 + c_2 - 2k_1^D + k_2^D} = \frac{a - 2c_2 + c_1 - 2k_2^D + k_1^D}{a - 2c_2 + c_1 - 2k_2^D + k_1^D + 3(c_2 - c_1) - 3(k_1^D - k_2^D)} < 1. \tag{C.33}$$

But, the right hand side of (C.32) is

$$\frac{a - 2c_1 + c_2 - 2(k_1^D + k_2^D)}{a - 2c_2 + c_1 - 2(k_1^D + k_2^D)} = \frac{a - 2c_1 + c_2 - 2(k_1^D + k_2^D)}{a - 2c_1 + c_2 - 2(k_1^D + k_2^D) - 3(c_2 - c_1)} > 1. \tag{C.34}$$

This is a contradiction. Therefore, we conclude that $k_1^D < k_2^D$. $\square$

Proof of Lemma 6

(i) Note that $m_i(-\infty, -\infty) = \infty$, $m_i(a - 2c_i + c_{-i}, a - 2c_i + c_{-i}) = 0$, and $m_i(k, k)$ monotonically declines with $k \in (-\infty, a - 2c_i + c_{-i}]$. By construction, $(m_i(k, k))^2$ is strictly decreasing in $k \in (-\infty, a - 2c_i + c_{-i}]$. Define $\overline{k}$ such that

$$(m_1^D)^2 + (m_2^D)^2 = (m_1(\overline{k}, \overline{k}))^2 + (m_2(\overline{k}, \overline{k}))^2. \tag{C.35}$$

If $\overline{k} \leq 0$, then $(m_1^D)^2 + (m_2^D)^2 \geq (m_1(0, 0))^2 + (m_2(0, 0))^2 > (m_1^U)^2 + (m_2^U)^2$. If $k^U > \overline{k} > 0$, then $(m_1^D)^2 + (m_2^D)^2 = (m_1(\overline{k}, \overline{k}))^2 + (m_2(\overline{k}, \overline{k}))^2 > (m_1^U)^2 + (m_2^U)^2$ because $m_i(\overline{k}, \overline{k}) > m_i^U$ for each $i \in \{1, 2\}$. Both cases imply that $e_1^D + e_2^D > e_1^U + e_2^U$. Finally, suppose to the contrary that $\overline{k} \geq k^U > 0$. From the supplier's optimal decision, we infer that

$$\pi_s(\overline{k}, \overline{k}) \leq \pi_s^D \Longrightarrow (m_1(\overline{k}, \overline{k}) + m_2(\overline{k}, \overline{k}))\overline{k} \leq m_1^D k_1^D + m_2^D k_2^D. \tag{C.36}$$

When $k^U \leq \overline{k}$, the first order condition at $k^U = \overline{k}$ implies that

$$-2(m_1(\overline{k},\overline{k}) + m_2(\overline{k},\overline{k}))(m_1(\overline{k},\overline{k}) + m_2(\overline{k},\overline{k}))\overline{k} + ((m_1(\overline{k},\overline{k}))^2 + (m_2(\overline{k},\overline{k}))^2)(3m_1(\overline{k},\overline{k}) + 3m_2(\overline{k},\overline{k}) - 2\overline{k}) \leq 0. \quad \text{(C.37)}$$

Rewriting (C.37) yields

$$\begin{aligned}\frac{2(m_1(\overline{k},\overline{k}) + m_2(\overline{k},\overline{k}))}{3m_1(\overline{k},\overline{k}) + 3m_2(\overline{k},\overline{k}) - 2\overline{k}} &\geq \frac{(m_1(\overline{k},\overline{k}))^2 + (m_2(\overline{k},\overline{k}))^2}{\overline{k}m_1(\overline{k},\overline{k}) + \overline{k}m_1(\overline{k},\overline{k})} \\ &\geq \frac{(m_1^D)^2 + (m_2^D)^2}{m_1^D k_1^D + m_2^D k_2^D} = \frac{2(m_1^D + m_2^D)}{3m_1^D + 3m_2^D - k_1^D - k_2^D}.\end{aligned} \quad \text{(C.38)}$$

The second inequality comes from (C.35) and (C.36). The last equality of (C.38) comes from the first order conditions of the supplier's profit maximization problem. However, since $2\overline{k} < k_1^D + k_2^D$ by construction,

$$\begin{aligned}\frac{2(m_1(\overline{k},\overline{k}) + m_2(\overline{k},\overline{k}))}{3m_1(\overline{k},\overline{k}) + 3m_2(\overline{k},\overline{k}) - 2\overline{k}} &= \frac{2a - c_1 - c_2 - 2\overline{k}}{3(2a - c_1 - c_2 - 4\overline{k})} \\ &< \frac{2a - c_1 - c_2 - (k_1^D + k_2^D)}{3(2a - c_1 - c_2 - 2(k_1^D + k_2^D))} = \frac{2(m_1^D + m_2^D)}{3m_1^D + 3m_2^D - k_1^D - k_2^D},\end{aligned} \quad \text{(C.39)}$$

which is a contradiction. The second strict inequality comes from condition $2\overline{k} < k_1^D + k_2^D$. Therefore, we conclude that $k^U > \max\{\overline{k}, 0\}$, $(m_1^D)^2 + (m_2^D)^2 > (m_1^U)^2 + (m_2^U)^2$, and $e_1^D + e_2^D > e_1^U + e_2^U$. Lemma 5, together with the above argument, says that $m_2^D < m_2(\overline{k},\overline{k}) < m_1(\overline{k},\overline{k}) < m_1^D$ because $(m_1(\overline{k},\overline{k}) - m_2(\overline{k},\overline{k})) < (m_1^D - m_2^D)$. Since $k^U > \overline{k}$, we obtain that $m_1^U < m_1(\overline{k},\overline{k}) < m_1^D$ and $e_1^U < e_1^D$.

(ii) The market share of seller 1 is given by

$$\frac{q_1^D}{q_1^D + q_2^D} = \frac{m_1(k_1^D, k_2^D)}{m_1(k_1^D, k_2^D) + m_2(k_2^D, k_1^D)} \quad \text{and} \quad \frac{q_1^U}{q_1^U + q_2^U} = \frac{m_1(k^U, k^U)}{m_1(k^U, k^U) + m_2(k^U, k^U)}. \quad \text{(C.40)}$$

Since $m_1(k_1^D, k_2^D) > m_1(k^U, k^U)$ and $m_2(k_2^D, k_1^D) < m_2(k^U, k^U)$,

$$\frac{m_1(k_1^D, k_2^D)}{m_1(k_1^D, k_2^D) + m_2(k_2^D, k_1^D)} > \frac{m_1(k^U, k^U)}{m_1(k^U, k^U) + m_2(k_2^D, k_1^D)} > \frac{m_1(k^U, k^U)}{m_1(k^U, k^U) + m_2(k^U, k^U)}. \quad \text{(C.41)}$$

Therefore, allowing third-degree price discrimination increases the market share of the cost-efficient seller and decreases that of the inefficient seller.

(iii) The ratio of seller 1's quantity under the discriminatory pricing rule to it under the uniform pricing rule is given by

$$\frac{q_1^D}{q_1^U} = \frac{(1 + (m_2(k_2^D, k_1^D)/m_1(k_1^D, k_2^D))^2)m_1(k_1^D, k_2^D)}{((m_1(k^U, k^U)/m_1(k_1^D, k_2^D))^2 + (m_1(k^U, k^U)/m_1(k_1^D, k_2^D))^2)m_1(k^U, k^U)} > 1, \quad \text{(C.42)}$$

since $m_1(k_1^D, k_2^D) > m_2(k_2^D, k_1^D)$ and $m_1(k_1^D, k_2^D) > m_1(k^U, k^U)$. This implies that when η is sufficiently close to one, $q_1^D > q_1^U$. $\square$

Proof of Lemma 7 Consider that the supplier chooses $(\tilde{e}_1^D, \tilde{e}_2^D, \tilde{k}_1^D, \tilde{k}_2^D)$ such that

$$\max_{e_1,e_2,k_1,k_2} k_1 \frac{(e_1 + e_2)}{b} m_1(k_1, k_2) + k_2 \frac{(e_1 + e_2)}{b} m_2(k_2, k_1) - \frac{1}{\eta} e_1^\eta - \frac{1}{\eta} e_2^\eta. \quad \text{(C.43)}$$

The first-order conditions are given by

$$e_1 \ : \ \frac{1}{b}(\tilde{k}_1^D m_1(\tilde{k}_1^D, \tilde{k}_2^D) + \tilde{k}_2^D m_2(\tilde{k}_2^D, \tilde{k}_1^D)) - (\tilde{e}_1^D)^{\eta-1} = 0, \quad \text{(C.44)}$$

$$e_2 \ : \ \frac{1}{b}(\tilde{k}_1^D m_1(\tilde{k}_1^D, \tilde{k}_2^D) + \tilde{k}_2^D m_2(\tilde{k}_2^D, \tilde{k}_1^D)) - (\tilde{e}_2^D)^{\eta-1} = 0, \quad \text{(C.45)}$$

$$k_1 \ : \ \frac{(\tilde{e}_1^D + \tilde{e}_2^D)}{b}\big(m_1(\tilde{k}_1^D, \tilde{k}_2^D) - \frac{2}{3}\tilde{k}_1^D + \frac{1}{3}\tilde{k}_2^D\big) = 0, \quad \text{and} \quad \text{(C.46)}$$

$$k_2 \ : \ \frac{(\tilde{e}_1^D + \tilde{e}_2^D)}{b}\big(m_2(\tilde{k}_2^D, \tilde{k}_1^D) - \frac{2}{3}\tilde{k}_2^D + \frac{1}{3}\tilde{k}_1^D\big) = 0. \quad \text{(C.47)}$$

Combining (C.44)-(C.47) yields

$$\tilde{e}_1^D = \tilde{e}_2^D = \Big[\frac{2}{b}\big(m_1^2(\tilde{k}_1^D, \tilde{k}_2^D) + m_1(\tilde{k}_1^D, \tilde{k}_2^D)m_2(\tilde{k}_2^D, \tilde{k}_1^D) + m_2^2(\tilde{k}_2^D, \tilde{k}_1^D)\big)\Big]^{\frac{1}{\eta-1}}. \quad \text{(C.48)}$$

From (C.46) and (C.47), we get the following optimal input prices

$$\tilde{k}_1^D = \frac{(a - c_1)}{2}, \quad \text{and} \quad \tilde{k}_2^D = \frac{(a - c_2)}{2}. \quad \text{(C.49)}$$

Since $c_1 < c_2$, we get $\tilde{k}_1^D > \tilde{k}_2^D$. Consider the uniform pricing rule in which the supplier solves for the same profit-maximizing problem (C.43) with one additional constraint, $k_1 = k_2 = k^U$. In this case, the optimal level of advertisement and input price are given by

$$\tilde{e}_1^U = \tilde{e}_2^U = \Big[\frac{3}{2b}\big(m_1(\tilde{k}^U, \tilde{k}^U) + m_2(\tilde{k}^U, \tilde{k}^U)\big)^2\Big]^{\frac{1}{\eta-1}}, \quad \text{and} \quad \tilde{k}^U = \frac{2a - c_1 - c_2}{4}. \quad \text{(C.50)}$$

By comparing $\tilde{k}^U$ with $(\tilde{k}_1^D, \tilde{k}_2^D)$, we get $\tilde{k}_2^D < \tilde{k}^U < \tilde{k}_1^D$. Given $\{\tilde{k}_2^D, \tilde{k}_1^D, \tilde{k}^U\}$, it is clear that $\tilde{e}_1^D = \tilde{e}_2^D > \tilde{e}_1^U = \tilde{e}_2^U$.

Recall that the aggregate output is given by

$$\tilde{q}_1^D + \tilde{q}_2^D = \frac{2\tilde{e}_1^D}{b}(m_1(\tilde{k}_1^D, \tilde{k}_2^D) + m_2(\tilde{k}_2^D, \tilde{k}_1^D)). \tag{C.51}$$

Since $m_1(\tilde{k}_1^D, \tilde{k}_2^D) + m_2(\tilde{k}_2^D, \tilde{k}_1^D) = m_1(\tilde{k}^U, \tilde{k}^U) + m_2(\tilde{k}^U, \tilde{k}^U)$ and $\tilde{e}_1^D + \tilde{e}_2^D > \tilde{e}_1^U + \tilde{e}_2^U$, we obtain that $\tilde{q}_1^D + \tilde{q}_2^D > \tilde{q}_1^U + \tilde{q}_2^U$. $\square$

BIBLIOGRAPHY

Acemoglu, D., "Training and innovation in an imperfect labour market," *The Review of Economic Studies*, 1997, *64* (3), 445–464.

Acemoglu, Daron and Jorn-Steffen Pischke, "Why Do Firms Train? Theory And Evidence," *The Quarterly Journal of Economics*, February 1998, *113* (1), 78–118.

Arya, Anil and Brian Mittendorf, "Input Price Discrimination When Buyers Operate in Multiple Markets," *The Journal of Industrial Economics*, 2010, *58* (4), 846–867.

Bagger, Jesper, Franois Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin, "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," *American Economic Review*, 2014, *104* (6), 1551–96.

Barro, Robert J and Jong Wha Lee, "A new data set of educational attainment in the world, 1950–2010," *Journal of development economics*, 2013, *104*, 184–198.

Bekman, Eli, John Bound, and Stephen Machin, "Implications of skill-biased technological change: international evidence," *The Quarterly Journal of Economics*, 1998, *113* (4), 1245–1279.

Bils, Mark and Peter J. Klenow, "Does Schooling Cause Growth?," *American Economic Review*, 2000, *90* (5), 1160–1183.

Bishop, John, "Is the Test Score Decline Responsible for the Productivity Growth Decline?," *The American Economic Review*, 1989, *79*, 178–197.

Bowlus, Audra J and Chris Robinson, "Human capital prices, productivity, and growth," *The American Economic Review*, 2012, *102* (7), 3483–3515.

Bulow, Jeremy I, John D Geanakoplos, and Paul D Klemperer, "Multimarket oligopoly: Strategic substitutes and complements," *Journal of Political economy*, 1985, *93* (3), 488–511.

Burdett, Kenneth and Dale T Mortensen, "Wage differentials, employer size, and unemployment," *International Economic Review*, 1998, pp. 257–273.

_ , **Carlos Carrillo-Tudela, and Melvyn G. Coles**, "Human Capital Accumulation And Labor Market Equilibrium," *International Economic Review*, 08 2011, *52* (3), 657–677.

Carneiro, Pedro and Sokbae Lee, "Trends in Quality Adjusted Skill Premia in the US, 1960-2000," *American Economic Review*, 2011, *101* (6), 2309–2349.

Charlot, Olivier and Bruno Decreuse, "Self-selection in education with matching frictions," *Labour Economics*, 2005, *12* (2), 251 – 267.

_ **and** _ , "Over-education for the rich, under-education for the poor: a search-theoretic microfoundation," *Labour Economics*, 2010, *17* (6), 886–896.

Choi, Jay Pil, "Optimal tariffs and the choice of technology Discriminatory tariffs vs. the Most Favored Nationclause," *Journal of International Economics*, 1995, *38* (1), 143–160.

Christensen, Bent Jesper, Rasmus Lentz, Dale T. Mortensen, George Neumann, and Axel Werwatz, "On the job search and the wage distribution," *Journal of Labor Economics*, 2005, *23* (1), 31–58.

Coase, R, "The Problem of Social Cost," *Journal of Law and Economics*, 1960, *3*, 1–44.

Córdoba, Juan Carlos and Marla Ripoll, "What explains schooling differences across countries?," *Journal of Monetary Economics*, 2013, *60* (2), 184–202.

DeGraba, Patrick, "Input market price discrimination and the choice of technology," *The American Economic Review*, 1990, *80* (5), 1246–1253.

Dertwinkel-Kalt, Markus, Justus Haucap, and Christian Wey, "Procompetitive dual pricing," *European Journal of Law and Economics*, 2015, pp. 1–21.

_ **,** _ **, and** _ , "Raising rivals cost through buyer power," *Economics Letters*, 2015, *126*, 181–184.

Flinn, Christopher and Joseph Mullins, "Labor market search and schooling investment," *International Economic Review*, 2015, *56* (2), 359–398.

Flynn, James R, "The mean IQ of Americans: Massive gains 1932 to 1978.," *Psychological bulletin*, 1984, *95* (1), 29.

Flynn, James Robert, *What is intelligence?: Beyond the Flynn effect*, Cambridge University Press, 2007.

Goldin, Claudia and Lawrence F. Katz, *The Race Between Education and Technology*, The Belknap Press of Harvard University Press, 2008.

Heckman, James J, Lance Lochner, and Christopher Taber, "Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents," *Review of economic dynamics*, 1998, *1* (1), 1–58.

_ , _ , **and** _ , "Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents," *Review of economic dynamics*, 1998, *1* (1), 1–58.

Hendricks, Lutz and Todd Schoellman, "Student Abilities During the Expansion of US Education," *Journal of Monetary Economics*, 2014, *63*, 19–36.

Herweg, Fabian and Daniel Müller, "Price discrimination in input markets: Downstream entry and efficiency," *Journal of Economics & Management Strategy*, 2012, *21* (3), 773–799.

Hobijn, Bart and Ayşegül Şahin, "Job-finding and separation rates in the OECD," *Economics Letters*, 2009, *104* (3), 107–111.

Hornstein, Andreas, Per Krusell, and Giovanni L Violante, "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *The American Economic Review*, 2011, pp. 2873–2898.

Inderst, Roman and Greg Shaffer, "Market power, price discrimination, and allocative efficiency in intermediate-goods markets," *The RAND Journal of Economics*, 2009, *40* (4), 658–672.

_ **and Tommaso Valletti**, "Price discrimination in input markets," *The RAND Journal of Economics*, 2009, *40* (1), 1–19.

Ishimaru, Shoya, Mai Seki, and Seung-Gyu Sim, "Degree Inflation and Hierarchical Labor Demand," 2015. Manuscript, August 2015.

Juhn, Chinhui, Dae Il Kim, and Francis Vella, "The expansion of college education in the United States: Is there evidence of declining cohort quality?," *Economic Inquiry*, 2005, *43* (2), 303–315.

Katz, Lawrence F and H David, "Changes in the wage structure and earnings inequality," in "Handbook of Labor Economics," Vol. 3A, Elsevier, 1999, pp. 1463–1555.

Katz, Lawrence F. and Kevin M. Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, February 1992, *107* (1), 35–78.

Katz, Michael L, "The welfare effects of third-degree price discrimination in intermediate good markets," *The American Economic Review*, 1987, pp. 154–167.

Kaymak, Barış, "Ability Bias and the Rising Education Premium in the United States: A Cohort-Based Analysis," *Journal of Human Capital*, 2009, *3* (3), 224–267.

Kim, Hyunchul and Seung-Gyu Sim, "Price discrimination and sequential contracting in monopolistic input markets," *Economics Letters*, 2015, *128*, 39–42.

Laitner, John, "Earnings within education groups and overall productivity growth," *Journal of Political Economy*, 2000, *108* (4), 807–832.

Lee, Chong Jae, "A Korean model of educational development," Technical Report, Center for Educational Research, Seoul National University 1992.

Lee, Hwa Ryung, Andras Pechy, and Michelle Sovinsky, "Is Intelrs Marketing Campaign Predatory?," 2013.

Levitt, Steven D, "Optimal incentive schemes when only the agents'" best" output matters to the principal," *The RAND Journal of Economics*, 1995, pp. 744–760.

Machin, Stephen and John Van Reenen, "Technology and changes in skill structure: evidence from seven OECD countries," *The Quarterly Journal of Economics*, 1998, *113* (4), 1215–1244.

Manuelli, Rodolfo E. and Ananth Seshadri, "Explaining International Fertility Differences," *The Quarterly Journal of Economics*, 2009, *124* (2), 771–807.

McAfee, R Preston and Marius Schwartz, "Opportunism in multilateral vertical contracting: Nondiscrimination, exclusivity, and uniformity," *The American Economic Review*, 1994, pp. 210–230.

Restuccia, Diego and Guillaume Vandenbroucke, "The evolution of education: A macroeconomic analysis," *International Economic Review*, 2013, *54* (3), 915–936.

_ **and** _ , "Explaining Educational Attainment across Countries and over Time," 2013. mimeo, University of Toronto.

_ **and** _ , "Explaining educational attainment across countries and over time," *Review of Economic Dynamics*, 2014, *17* (4), 824 – 841.

Rey, Patrick and Thibaud Vergé, "Bilateral control with vertical contracts," *RAND Journal of Economics*, 2004, pp. 728–746.

Salant, Stephen W and Greg Shaffer, "Unequal treatment of identical agents in Cournot equilibrium," *American Economic Review*, 1999, pp. 585–604.

Shimer, Robert, "On-the-job search and strategic bargaining," *European Economic Review*, 2006, *50* (4), 811 – 830.

Sim, Seung-Gyu, "Wage Dynamics with Private Learning-by-doing and On-the-job Search," September 2014. Manuscript, September 2014.

_ , "On-the-Job Training and On-the-Job Search: Wage-Training Contracts in a Frictional Labor Market," August 2015. Manuscript, August 2015.

Stewart, Jay, *Recent trends in job stability and job security: Evidence from the March CPS*, US Department of Labor, Bureau of Labor Statistics, Office of Employment and Unemployment Statistics, 2002.

Taubman, Paul and Terence Wales, *Mental Ability and Higher Educational Attainment in the Twentieth Century*, National Bureau of Economic Research, 1972.

Topel, Robert and Michael Wald, "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, May 1992, *107* (2), 439–479.

Valletti, Tommaso M, "Input price discrimination with downstream Cournot competitors," *International Journal of Industrial Organization*, 2003, *21* (7), 969–988.

Yoshida, Yoshihiro, "Third-degree price discrimination in input markets: output and welfare," *The American Economic Review*, 2000, *90* (1), 240–246.