ESSAYS IN THE ECONOMICS OF EDUCATION AND EXPERIMENTAL ECONOMICS

by

Gumilang Aryo Sahadewo

S.E. in Economics, Universitas Gadjah Mada, 2009

M.A. in Economics, University of Pittsburgh, 2013

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 PhD in Economics

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Gumilang Aryo Sahadewo

It was defended on

May 31, 2017

and approved by

Daniele Coen-Pirani, PhD, Associate Professor

Daniel M. Berkowitz, PhD, Professor

Marla Ripoll, PhD, Associate Professor

Sera Linardi, PhD, Assistant Professor

Lindsay C. Page, PhD, Assistant Professor

Dissertation Director: Daniele Coen-Pirani, PhD, Associate Professor

ESSAYS IN THE ECONOMICS OF EDUCATION AND EXPERIMENTAL ECONOMICS

Gumilang Aryo Sahadewo, PhD

University of Pittsburgh, 2017

This dissertation consists of three studies in the economics of education and experimental economics. In Chapter 1, I address a debate in the literature about the effects of measures of school quality on labor market earnings. Using individual-level data, previous studies find no effects of measures of school quality and a subsequent study argues that the result is driven by the sample that includes mainly young individuals. I use recent NLSY79 Geocode data that provides extended earnings observations including prime-age earnings. I find that the percentage of teachers with a Master's degree has a positive long-run effect on individuals' earnings in the labor market. In Chapter 2, we examine the Dell Scholars Program which provides a combination of financial support and individualized advising to selected students throughout their postsecondary experience. We capitalize on an arbitrary cutoff in the program's algorithmic selection process and a regression-discontinuity analytic strategy. We find that, at the margin of eligibility, being selected as a Dell Scholar has positive impacts on later persistence and on-time bachelor's degree completion. Finally, in Chapter 3, we conduct a series of laboratory experiments to explore the effects of religion on prosocial risk taking. We find that the religious message can induce prosocial risk taking only when doing so help others of the same beliefs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. SCHOOL QUALITY AND LABOR MARKET EARNINGS: SOME NEW RESULTS ON AN OLD DEBATE

1.1. INTRODUCTION

There is an open debate about the effects of measures of school quality on labor market earnings. Betts (1995) investigates the direct effects of measures of public high-school quality—teacher-student ratio, relative teacher salaries, and the percentage of teachers with a Master's degree—on individuals' earnings between 1979 and 1990 using the 1979 National Longitudinal Survey of Youth (NLSY79) data. He finds no direct effects of the measures of school quality on individuals' earnings. Card and Krueger (1996) argue that Betts' sample is quite young as the average age of individuals in the sample is only 23.8 years old. They claim that the effects of the measures of school quality on earnings are understated among young individuals. Individuals who attended better-quality high schools might have acquired a higher level of education and were absent from the labor market. Consequently, Betts' sample underrepresented individuals with potentially higher earnings in the labor market.

This paper addresses this issue by analyzing the effects of measures of school quality on earnings using more recent data from the NLSY79. The data provide extended earnings observations which allow us to analyze the effects of measures of school quality when individuals were older and reduce the issue of selection into schooling considerably. Moreover, the extended earning observations convey better information about individuals' labor market potentials. I collect an unbalanced panel data using a public-version of the NLSY79 data, that includes earnings observations from 1979 to 2010, and the Geocode version of the NLSY79 data. For comparability with Betts' sample, I proceed in three steps. First, I replicate Betts' sample restriction procedure to approximate his 1979-1990 sample as closely as possible. Second, I use the recent NLSY79 to obtain post-1990 data on earnings, the number of weeks worked, and other individuals' characteristics. Third, I conduct regression analysis using Betts' main specification to investigate the effects of measures of school quality on earnings when individuals were older.

Even after closely following Betts' sample restriction procedure, my sample has more individuals than Betts' sample. Retrospective revisions of the NLSY79 data post-1990 explain this difference. Nevertheless, I replicate Betts' findings that the measures of school quality do not affect individuals' earnings in the analysis of the 1979-1990 sample. However, I find that the percentage of teachers with a Master's degree has a positive and significant effect on individuals' earnings when they were 40 or older. Moreover, the effects of the percentage of teachers with a Master's degree on individuals' earnings are increasing over individuals' life cycle. I find that the effect of the percentage of teachers with a Master's degree is concentrated among high school graduates. These results are not sensitive to additions of important covariates such as parents' education, family wealth, and individuals' age-adjusted AFQT scores. These findings are important as they suggest that selection into better quality schools is not an issue. Overall, the findings in this paper support the claim that measures of school quality affect earnings only after individuals were older.

Findings in this study are quite important as they reconcile two strands of studies in the literature that fail to reach a consensus regarding the effects of school quality on earnings. In one strand, researchers use measures of school quality observed at the state or aggregate level, such as per-student expenditure, teacher-student ratio, and state average teacher salary (Wachtel, 1976; Rizzuto and Wachtel, 1980; Card and Krueger, 1992b,a; Nechyba, 1990). There are two main advantages of using state or aggregate-level measures. First, the measures average out endogenous measures of school quality. The endogeneity arises because families select into different schools based on observable and, more importantly, unobservable characteristics. Second, the measures average out errors (Card and Krueger, 1994, 1996). Overall, state-level studies find significant effects of measures of school quality on earnings with elasticities ranging from 0.08 to 0.16.

The other strand of the literature investigates the effects on individual earnings of school-level measures of school quality (Betts, 1995; Grogger, 1996; Betts, 2001; Strayer, 2002; Tobias and Li, 2003).[1] Betts (1995) argues that an analysis using school-level measures is superior as it eliminates aggregation bias in studies using state or aggregate-level measures of school quality.[2] He then estimates the effects of measures of school quality, such as the student-teacher ratio, the percentage of teachers with a Master's degree, and the relative teacher salaries, on individuals' earnings. Using the 1979-1990 NLSY79, Betts collects a sample of individuals aged 32 or younger for his estimations. Betts (1995) finds that the measures of school quality at the school level do not affect individuals' earnings. Strayer (2002) who uses a similar sample from the NLSY79 also concludes that there is no effect of the measures of school quality on earnings.

A study by Grogger (1996) analyzes whether the black-white wage gap trend could be attributed to differences in school quality. Using the 1972 National Longitudinal Study (NLS72) and the High School and Beyond (HSB) survey, Grogger (1996) estimates weak relationships between the measures of school quality and individuals' earnings. He finds that the measures of school quality, such as student to teacher ratio, and term length were not a good predictor of earnings. He does find that teacher education, measured by a dummy variable that equals to one if individuals attended schools in which 30 percent or

[1]A related study in the literature is by Dearden et al. (2002) who investigate the effects of school-level measures of school quality on earnings in the UK.

[2]Consider an estimation of school quality effects using state-level measures of school quality. This measure captures variation of school quality between states. However, it does not capture variation within each state which could be substantial. Thus, using state averages instead of school-level data eliminates the within-state variation, which may lead to bias school quality effects. See Sellin (1990) for a detailed discussion on aggregation bias.

more teachers have advanced degrees, has a significant effect on earnings for individuals in the NLS72 sample. It is important to note that this result can be quite sensitive to changes in the definition of the teacher education variable. He acknowledges that the sample was relatively young. He re-estimates the effects of the measures of school quality when the individuals in the NLS72 were about 32 years old. Still, he finds similar results.

Betts (1996) addresses Card and Krueger's criticism by investigating the effects of the measures of school quality on individuals' earnings only after the individuals acquired a considerable amount of on-the-job experience. Betts (1996) uses earnings of individuals aged 40-55 in the 1980 Census to predict NLSY79 individuals' earnings when they were 40-55 years old. Betts concludes that the measures of school quality do not have significant effects on individuals' prime age earnings. This study has a couple of caveats. First, the 1980 Census is cross-sectional data, hence we can only observe one earnings observation for each individual. One needs several different individuals to construct a series of earnings observations for ages 40-55. However, each individual has different observed as well as unobserved characteristics, which may significantly affect measured earnings in different ways. Second, the earnings obtained from the 1980 Census cannot capture the actual state of the economy and technological progress, which significantly affect occupation-specific productivity and earnings growth. Therefore, earnings data for a specific occupation in the 1980 Census are not comparable to those post-1990.

This study is also related to a recent study by Chetty et al. (2011) who estimate the effects of measures of school quality, more specifically classroom quality, on earnings. Chetty et al. (2011) link the original STAR project data to administrative tax-returns data. They found no significant impact of smaller class sizes on earnings. Nevertheless, their estimates suggest that students who were taught by more experienced teachers earned higher than students who were taught by less experienced teachers. Lastly, they estimate the effects of classroom quality proxied by average test scores of a student's classmates on earnings. They find a significant effect of class quality on the individuals' earnings in the labor market.

This study is also related to studies that investigate the effects of school quality on earnings in developing countries and the effects of school quality on return to education among immigrants. A study by Bedi and Edwards (2002) examines the effects of measures of school quality—such as teacher training, school infrastructure, and measures of school crowding—on earnings in Honduras using micro data. They find positive and significant effects of the measures of school quality on individuals' earnings. A study by Bratsberg and Terrell (2002) investigates the effects of measures of school quality on US return to education of immigrants. They find negative and significant effects of student-teacher ratio on US return to education.

The rest of the chapter is organized as follows. In Section 1.2, I discuss Card and Krueger's criticism that the effects of measures of school quality on earnings are understated among young individuals. In section 1.3, I discuss the NLSY79 dataset used in the regression analyses. I also discuss in detail a replication of Betts' (1995) sample selection procedures, and I compare descriptive statistics of my sample to those in Betts (1995). Section 1.4 discusses the regression specifications and the estimation results. I state my conclusion in section 1.5.

1.2. CARD AND KRUEGER'S CRITICISMS

Card and Krueger (1996) argue that the sample used in studies using school-level measures, for instance Betts (1995) and Grogger (1996), is quite young. Betts (1995) uses the NLSY79 data and his sample include individuals who were 17 to 32 years old with an average age of 23. Grogger (1996) uses the NLS72 and the HSB data which include individuals who were 23-25 years old. Card and Krueger (1996) claim that a relatively young sample understates the effects of the measures of school quality on earnings. In this section, I discuss two main arguments that Card and Krueger use to support their claim.

First, earnings observations of young individuals may not convey a complete information about individuals' earnings potential in the labor market (Card and Krueger, 1996). Many young individuals were not in the labor market because they attended school. Even if they were in the labor market, many of them worked part time. I verify these arguments by analyzing schooling and employment patterns among young individuals in the NLSY79 sample.

I identify in the NLSY79 individuals who attended school and those who worked. I also identify whether individuals worked part or full-time jobs using cumulative hours worked. Individuals worked part-time jobs if they worked less than 2080 hours while individuals worked full-time jobs if they worked more than 2079 hours. As shown in Figure 1.1, there was a considerable share of young individuals in the 1979-1990 sample who attended school. Between 1979 and 1982, more than 10 percent of individuals aged 17 or older were enrolled in school annually. I find that about 96 percent of individuals who attended school worked part-time jobs. I also show that there was also a considerable share of young individuals who worked part time in Betts' sample. For example, more than 40 percent of the individuals worked part time between 1979 and 1982.

Most of the young individuals were also at an early stage of their careers during which young individuals might have spent quite amount of time for job trainings or might have had a higher tendency to change job. An empirical study by Lynch (1992) shows that young male individuals in the NLSY79 sample who took job trainings spent 34 to 74 weeks to attend job trainings. Topel and Ward (1988) find that 80 percent of young individuals with a decade of working experience changed jobs at least three times and that more than half of similar young individuals changed jobs more than 6 times.

Second, OLS estimates of the effects of the measures of school quality are biased downward due to a selection issue. We face this issue because we can only observe earnings for individuals who worked part or full-time jobs. The intuition that the effects of the measures of school quality are biased downward is straightforward. Individuals who attended schools with lower measures of quality are less likely to enroll in a college and are more likely to

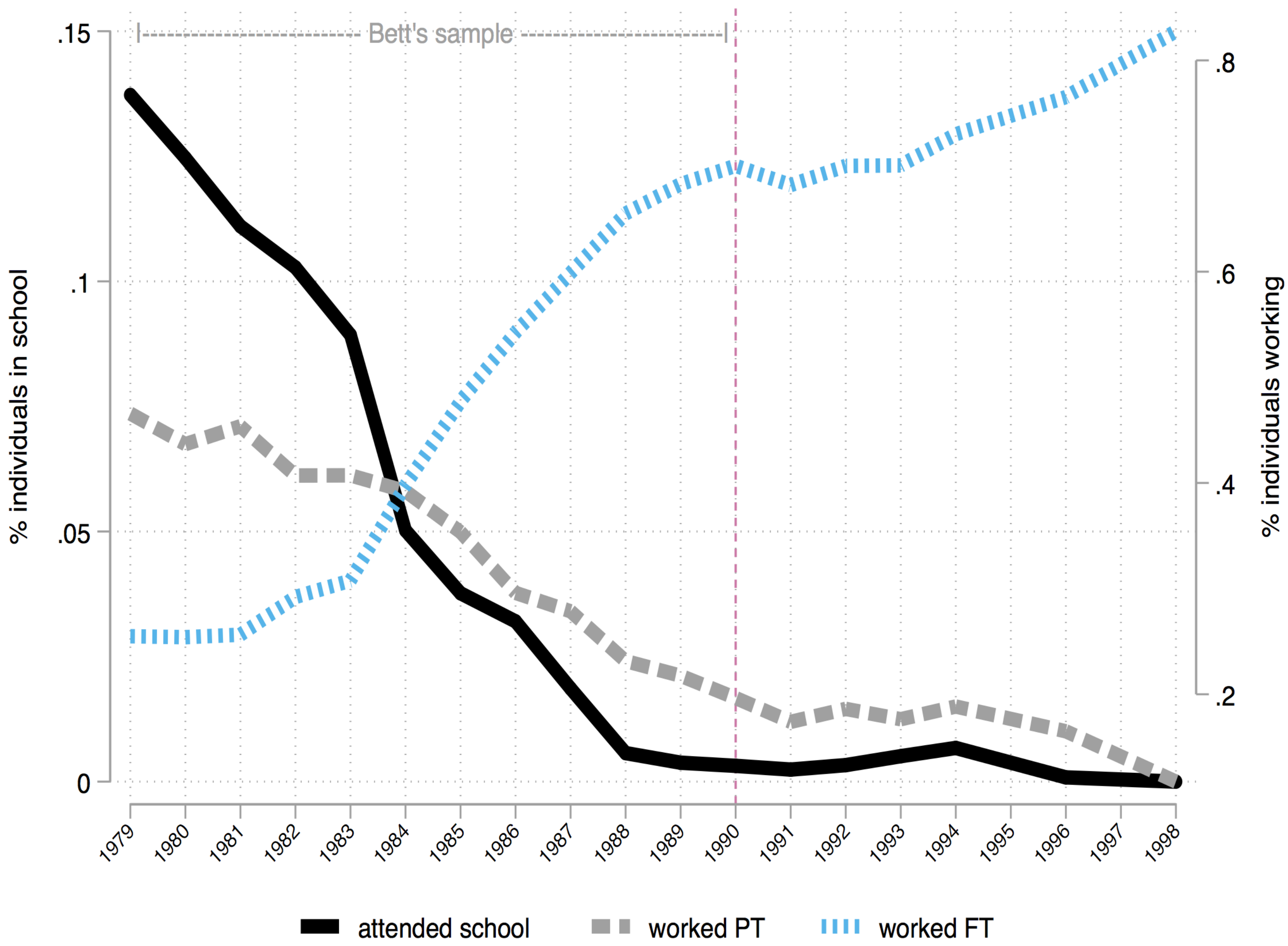

Figure 1.1: Schooling and Employment Pattern, 1979-1994. Source: Author's calculation using the NLSY79.

participate in the labor market. On the other hand, individuals who attended schools with higher measures of quality are more likely to enroll in a college, worked part-time jobs, and earned relatively lower earnings. Thus, the effects of the measures of school quality are biased downward.

The discussions above suggest that many young individuals were enrolled in school or worked part-time jobs. A logical follow-up inquiry is: are measures of school quality correlated with schooling and employment status among young individuals in the NLSY79 sample? I estimate a logit model to analyze the correlations between measures of school quality and employment status of individual i at time t:

$$P\left(y_{it}=1 \mid X_{it}\right)=G\left(\beta_0+\theta Quality_i+\Gamma \mathbf{X}_i\right) \tag{1.1}$$

where y is either attending school (while working), working part-time jobs, or working full-time job. The vector $\mathbf{X}$ includes individual-specific characteristics, family characteristics, time dummies, and region dummies. I estimate the model using the 1979-1990 NLSY79 sample, which is comparable to Betts' sample, and I cluster the standard errors at the individual level. I report the results of the estimations in Table 1.1.

There is suggestive evidence that the measures of school quality are positively correlated with workers' decisions to attend school. In Column 1 of Table 1.1, we can observe that the percentage of teacher with Master's degree and the teacher-student ratio associated with a higher probability of attending school. The correlation between the percentage of teacher with Master's degree and schooling decisions is in line with the main findings in Strayer (2002). Using NLSY79 data, he shows that a higher percentage of teachers with a Master's degree is associated with a higher probability of college enrollment. However, the correlations between these variables and schooling decisions are no longer significant when I include additional covariates such as father's education and AFQT test score into the model specification (Column 4 of Table 1.1). This is due to a strong correlation between these

Table 1.1: Correlations between Measures of School Quality and Schooling Decisions

	1: 1 if in school	2: 1 if working PT	3: 1 if working FT	4: 1 if in school	5: 1 if working PT	6: 1 if working FT
teacher with a Master's	0.000392***	0.000307	-0.000264	0.0000713	0.000158	-0.000237
	(0.000135)	(0.000238)	(0.000325)	(0.000110)	(0.000238)	(0.000310)
teacher-student ratio	0.359**	-1.027***	1.196***	0.203	-0.995***	1.160***
	(0.161)	(0.348)	(0.459)	(0.136)	(0.340)	(0.428)
relative teacher salaries	0.00345	0.00190	0.000559	0.00148	0.000415	-0.00126
	(0.00251)	(0.00541)	(0.00677)	(0.00218)	(0.00551)	(0.00638)
Observations	14,859	15,589	15,589	14,859	15,589	15,589
Other Controls	N	N	N	Y	Y	Y
Cluster SE	Individual	Individual	Individual	Individual	Individual	Individual

Source: author's calculation using the NLSY79.

Notes: standard errors are clustered at the individual level and are shown in the parentheses. The signs *, **, *** indicates significance at 10, 5, and 1 percent. Other regression covariates not shown in the table are years of schooling, age, age-adjusted AFQT score, father's education, family's wealth in 1980, an SMSA dummy, a marriage dummy, region dummies, year dummies, interaction between year and region dummies, and time dummies. The sample include individuals aged 17-32 years old.

measures of school quality with AFQT score and father's education.

In Column 2 and 3 of Table 1.1, we can observe that the teacher-student ratio is correlated with the probability of working part and full-time jobs. A higher teacher-student ratio is correlated with a lower likelihood of working part-time jobs but it is correlated with a higher likelihood of working full-time jobs. The correlations persist even after I control for additional covariates (Column 5 and 6 of Table 1.1).

In the next section, I test Card and Krueger's claim that the measures of school quality affects earnings when individuals were older. I proceed in two steps. First, I investigate the effects of the measures of school quality when the individuals were 17-32 by replicating the estimations in Betts (1995). I also re-estimate Betts' model using clustered standard errors to accommodate the fact that individuals' unobserved characteristics are correlated across time. I then investigate the effects of the measures of school quality on prime-age earnings when the individuals were 40-50.

1.3. DATA

1.3.1 Sample Selection Procedures

The data for the analysis are obtained from the public version of the National Longitudinal Survey of Youth (NLSY79) and the Geocode version of the NLSY79. The data provide a representative sample of 12,686 young individuals and the data record individuals' information from 1979 to 2010. Annual interviews were conducted for these individuals through 1994 and biennial interviews were conducted in subsequent periods until 2010. The data provide information regarding labor market behaviors and outcomes, education history, family background, and government participation program. Using the Geocode version of the NLYS79, I obtain individuals' states of residence from 1979 to 2010 and states of schools that individuals attended. I use these variables to calculate relative salaries of teachers with a B.A. degree and in the regression specification to control for state-specific shocks. More importantly, the states of residence variable allows me to replicate Betts' estimation. I list the key variables used in the model specifications in Table A.1.

Using the NLSY79, I obtain the measures of school quality from a survey of high school attended by NLSY respondents conducted in 1979. The survey reported many school characteristics such as the number of students enrolled, the number of full-time teacher, the teachers' educational attainment, academic-related facilities, the curriculum and the school activities. The main measures of school quality used in the typical regression analyses are teacher-student ratio, the salaries of teachers with a B.A. degree, and the percentage of teachers with a Master's degree. I choose these measures for a direct comparability with the main analyses in Betts (1995) and Betts (1996).[3] I adjust the salaries of teachers with 1980 per-capita personal income in the state to calculate relative teacher salaries.[4]

I conduct a three-step sample restriction procedure to obtain the 1979-2010 sample. First, I follow Betts' sample selection procedure. Individuals with missing or declining years of schooling between 1979-1990 are excluded from the sample. Individuals in the military subsample, whose measures of school quality are not observed, those who did not attend a public high school, male Hispanics, male Blacks, and females are also excluded from the sample. Lastly, there were individuals who were not enrolled in a junior high school during the 1979 high-school survey. They were either very young, were currently enrolled in a junior high school at the time of survey, or did not attend a high school in 1979. I exclude these individuals since we do not observe the measures of school quality. From these restriction criteria, I obtain a sample of about 2,077 white-male individuals who attended public high schools in 1979.

[3]Betts (1995) also analyses the effects of the number of student enrollment, the percentage of disadvantaged students, the percentage of students dropping out, library books per enrollee, the availability of vocational curricula, the proportion of black students, and the percentage of teachers who had left in the previous year for reasons other than death and retirement. He finds a positive and significant correlation between enrollment and earnings. He also finds negative effects on earnings of the percentage of disadvantaged students and the percentage of students dropping out for individuals who did not complete high school.

[4]I obtain per-capita personal income data from Table SA1-3 in the BEA Regional Economic Accounts website: https://www.bea.gov/regional/.

Some individuals in this subsample have missing earnings and number of weeks worked observations. Some individuals also reported zero earnings in the survey. In these cases, these observation points, but not the individuals, are excluded from the sample. Some individuals are also missing observations on marital status, residence in SMSA, and state of residence. In these cases, I assume that individuals keep the same status as in the previous period and impute the missing observations using values from previous period.

Lastly, observations of individuals who lived outside of United States in any given year are also excluded from the sample. These restrictions reduce the 1979-1990 sample to an unbalanced panel consisting of 1,568 individuals. In this sample, the average number of wage observations per individuals is 10.2. For a comparison, the number of individuals in Betts' sample was 1,134.

Table 1.2: A Summary of the Sample Restriction Procedure for the 1979-1990 Sample

Procedures	**Betts (1995)**		**This study**	
	# of indiv.	**obs.**	**# of indiv.**	**obs.**
initial			12,669	304,056
drop ind. missing education			8,744	209,856
drop ind. in military subsample			8,617	206,808
drop obs. beyond 1990			8,617	103,404
drop obs younger than 17 y.o.			8,617	97,513
drop obs. miss. earnings & weeks worked	6,749	68,182	8,470	74,096
keep white males		19,534	2,350	22,325
keep public school students		18,395	2,200	20,913
keep those in US states & DC		17,706	2,200	20,912
keep positive weekly earnings		16,417	2,200	20,841
elim. student teacher ratio of 6		16,406	2,199	20,813
keep those w/ school quality		11,314	1,568	14,955

Source: author's calculation using the NLSY79.

Table 1.2 compares the number of individuals and observations in Betts' 1979-1990 sample and those in this study. Ideally, I want to obtain a sample identical to that used in Betts (1995). However, I obtain significantly a higher number of observations for the 1979-1990 sample. This difference is due to retrospective revisions or imputations by the surveyor to recover information lost in the previous survey years. Specifically, the information in BLS website states: "*from this information and other retrospective information, a longitudinal record spanning from the date of, and to some extent the time preceding, the first interview through the most current interview date can be constructed for each respondent. The longitudinal record is maintained even for respondents who are not interviewed in interim years. Each year's questionnaire incorporates retrospective questions designed to recover as completely as possible information lost (or incorrectly reported or recorded) during previous survey years.*"

Table 1.3: The Summary Statistics of the 1979-1990 Sample

Variable	**Betts (1995)**		**This study**			
			Initial Sample		**Regression Sample**	
	mean	**std.dev.**	**mean**	**std.dev.**	**mean**	**std.dev.**
log of weekly earnings	5.5229	0.8292	5.4036	0.903	5.3804	0.8800
education completed	12.4360	2.1295	12.7941	2.2552	12.8655	2.0558
experience (weeks)	216.1386	161.6807	248.0426	154.2158	230.2812	140.8724
marital status	0.3337	0.4715	0.3798	0.4852	0.3604	0.4801
residence SMSA	0.6801	0.4664	0.6947	0.4606	0.7022	0.4573
teacher-student ratio	0.0559	0.0155	0.05587	0.0155	0.0564	0.0150
teacher with master's	46.1171	22.7448	46.4558	22.8588	46.2022	23.0822
relative teacher salaries	0.85384	0.10262	1.2010	0.1589	1.1971	0.1654
# of Individuals	n.a.		1,568		1,038	

Source: author's calculation using the NLSY79.

I present the summary statistics of Betts' and my 1979-1990 samples in Table 1.3. I find that the averages and standard deviations of most of the variables are close to those in Betts (1995). There are a couple of noticeable differences. First, the average total number of weeks worked in our sample is higher by about 32 weeks in average. Second, the average relative salaries of teachers is also higher by 0.20 in our sample. I may have used different adjustment variable as the link in Betts (1995) is no longer available. Hanushek (1986) document measures of school quality such as the teacher-student ratio and the percentage of teachers between 1960 and 1980. I find that the averages of these measures calculated from the NLSY79 data are quite close to the ones documented in Hanushek (1986).

Second, I use the recent version of NLSY79 to obtain extended earnings observations between 1991 and 2010. Several variables have missing or declining values post 1990. I use the following procedures to impute many missing values of the explanatory variables in the subsequent observation periods. First, I assume that individuals did not enroll in school after 30. This is a quite fair assumption as the data suggests that only about 2.23 percent of individuals aged 30 or older were enrolled in school. These individuals were either enrolled in a college or a graduate program. I then impute missing or declining years of schooling of individuals aged 30 or older with values from the preceding period. Second, I exclude observation periods with missing earnings and number of weeks worked in each period. I obtain, in average, additional 10.6 periods of observations for each individual in the sample between 1991-2010. The total number of individuals in the unbalanced panel is 1,568 with 26,802 observations.

1.3.2 Correlations between the Measures of School Quality

There are several variables in the NLSY79 that can be used to measure school quality. I focus on the teacher-student ratio, the relative teacher salaries, and the percentage of teachers with a Master's degree which are commonly use as measures of school quality in the literature.

I also analyze the number of students enrolled, the percentage of economically disadvantaged students, the percentage of grade 10 students who dropped out, the percentage of teacher who left the school, the percentage of black students, and the number of books per student to gain a better understanding about the main measures of school quality. Heckman et al. (1996) suggest that it is important to understand how different measures of school quality are correlated with each other. They argue that it is possible for the measures of school quality to be negatively correlated, for example, due to budget constraints.

I report pairwise correlation coefficients between the measures of school quality in Table 1.4. There is a positive correlation between the percentage of teachers with a Master's degree and the relative salaries of teachers with a BA degree. This finding suggests that, on average, high schools who employed a higher percentage of teachers with a Master's degree also paid higher salaries to teachers with a B.A. degree. Thus, high schools might have increased the classroom size owing to limited budgets which leads to a negative correlation between the relative salaries and the teacher-student ratio. Schools with a higher percentage of teachers with a Master's degree tend to have a larger enrollment which leads to a negative correlation between the percentage of teachers with a Master's degree and the teacher-student ratio.

The results suggest that a higher percentage of teachers with a Master's degree seems to be an indicator of a better school quality. For instance, the percentage of teachers with a Master's degree is negatively correlated with the percentage of disadvantaged students. It is also negatively correlated with the percentage of grade 10 students who dropped out and the percentage of teacher who left the school. Overall, the results show that the teacher-student ratio, the relative teacher salaries, and the percentage of teachers with a Master's degree altogether do not indicate improvement in school quality. For this reason, I do not combine the main measures of school quality into a single index.

In the next section, I estimate regression models for different age ranges to analyze the effects of measures of school quality on earnings using the extended earning observations. First, I focus on the teacher-student ratio, the relative teacher salaries, and the percentage of teachers with a Master's degree as the main measures of school quality. The extended earning observations allow me to estimate the effects of the measures of school quality when individuals were older. Specifically, I estimate the model in two age ranges: when the individuals were 17-32 and when the individuals were 40 or older. These age ranges allow me to compare estimation results with those in Betts (1995) and Betts (1996).

For each age range, I keep the same group of individuals to avoid a composition effect. Thus, I exclude individuals from the sample if the individuals have no earnings observation in one or any of the two age ranges. The total number of unique individuals with earning observations in the two age ranges is 1,038. For a comparison, the number of individuals observed in each age range is 1,568 and 1,038 respectively. Regressions with different group of individuals yield similar results.

Table 1.4: Correlation Coefficients of the Measures of School Quality

	t-s ratio	salaries	master	enrollment	disadvantaged	dropout	teacher left	blacks	books
Teacher-student ratio	1								
Relative teacher salaries	-0.126	1							
% of teachers with a Master's	-0.104	0.359	1						
Total enrollment	-0.427	0.205	0.341	1					
% of disadvantaged	0.009	-0.183	-0.245	-0.187	1				
% of dropout	-0.025	0.007	-0.147	-0.024	0.254	1			
% of teachers who left	0.219	-0.204	-0.272	-0.178	-0.016	-0.022	1		
% of black students	-0.095	0.014	0.036	0.168	0.285	0.100	-0.042	1	
Library books per student	0.202	-0.151	-0.164	-0.262	0.126	0.053	0.192	-0.080	1

Source: author's calculation using the NLSY79.

1.4.1 The Effects of the Measures of School Quality on Earnings

The main objective of analysis in this section is to replicate Betts' estimation and to investigate the effects on earnings of the measures of school quality when individuals were older. I start by replicating Betts' estimations using the 1979-1990 NLSY79 sample, in which individuals were 17-32 years old. For the estimation, I use Betts' model specification:

$$\ln w_{it} = \alpha + \rho S_{it} + \theta \mathrm{Q}_i + \gamma \mathrm{E}_{it} + \beta \mathrm{X}_{it} + \epsilon_{it}. \tag{1.2}$$

where i indicates an individual in the sample at time t. As discussed in the previous section, I estimate this model using the sample of individuals who are observed in the two age ranges. Note that the data include earnings of individuals who worked. Thus, I acknowledge that my analyses may be subject to the sample selection issues (Heckman, 1979; Vella, 1998).

The left-hand side variable is a log of weekly earnings however results are similar in regressions with a log of hourly earnings. The regressors are the years of schooling (S), the measures of school quality (Q), a quartic function of experience (E), and a vector of control variables (X). The measure of labor market experience at a given period is the sum of the number of weeks worked in previous periods. The control variables include 8 dummy variables for residences in the census regions, interaction terms between the census region dummies and the years of schooling, a dummy variable for residence in SMSA, and a dummy for the marital status, and year dummies. Betts (1995) uses heteroscedasticity-robust standard errors estimators for pooled OLS regressions. However, the error terms ϵ_{it} are likely to be correlated over time for a given individual owing to the structure of the data. For this reason, I consider an additional baseline specification with standard errors clustered at the individual level (Cameron and Miller, 2015).

I first estimate the model using the 1979-1990 sample to compare the results with Betts'. I report my estimates and Betts' estimates in Table 1.5. I present Betts' estimates in Column 1 and I present my estimates in Column 2 and Column 3. Betts' estimates show that measures of school quality have no effect on individuals' earnings. My estimates in Column 2 show that the teacher-student ratio has a positive effect on individuals' earnings. However, the significance goes away once I cluster the standard errors at the individual level. In all model specification, the percentage of teacher with a Master's degree and the relative teacher salaries have no effect on individuals' earnings.

We can observe that the estimated standard errors are higher in the specification using the clustered standard errors. This increase is expected in time-ordered earnings observations as the error terms are positively correlated within individual level (Cameron and Miller, 2015). For this reason, I cluster the standard errors at the individual level in subsequent specifications to control for the correlations of the error terms across time. I acknowledge that the point estimates for the measures of school quality are different than those in Betts (1995). These differences are not unexpected because the regression sample in Betts' and the sample in this study are different. Nevertheless, I obtain similar qualitative results with Betts' that the measures of school quality do not affect individuals' earnings.

To analyze the effects of the measures of school quality on earnings when individuals were older, I estimate Specification 1.2 at different age ranges: 17-32 and beyond 40. I estimate the model using the same sample to avoid composition effects and I cluster the standard errors at the individual level. I present the estimates for the two age ranges in Column 3 and Column 5 of Table 1.5, respectively. The estimates in Column 3 of Table 1.5 show that no measures of school quality appear to be significantly different from zero when individuals were 17-32. However, the estimates in Column 5 of Table 1.5 show that a higher percentage of teachers with a Master's degree leads to significantly higher earnings when individuals were 40 or older. The point estimate suggests that a one standard-deviation increase in the percentage teachers with a Master's degree increases the prime-age earnings by 0.07 standard deviation. For a comparison, a one standard-deviation increase in the years

Table 1.5: The Estimated Effects of Measures of School Quality on Earnings

	Betts (1995)	Replication of Betts (1995)			Extended Observations	
	1: 17-32 y.o.	2: 17-32 y.o.	3: 17-32 y.o.	4: 17-32 y.o.	5: 40+ y.o.	6: 40+ y.o.
Education	0.041***	0.029***	0.029	0.080	0.120***	0.156**
	(0.012)	(0.012)	(0.025)	(0.056)	(0.021)	(0.072)
Teacher-student ratio	0.250	1.126**	1.126	1.149	2.051	2.112
	(0.439)	(0.525)	(1.077)	(1.085)	(1.305)	(1.317)
Relative teacher salaries	-0.039	-0.065	-0.065	-0.031	-0.111	-0.078
	(0.638)	(0.046)	(0.085)	(0.095)	(0.132)	(0.138)
% of teachers with a Master's	0.000	-0.000	-0.000	-0.000	0.00261***	0.00257***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Education · teacher-student ratio				-0.398		0.076
				(0.440)		(0.572)
Education · teacher salaries				-0.032		-0.031
				(0.034)		(0.052)
Education · teachers with a Master's				-0.000		-0.000
				(0.000)		(0.000)
Observations	11,314	9,190	9,190	9,190	3,146	3,146
Adjusted R^2	0.313	0.377	0.377	0.377	0.225	0.225
Number of individuals	1,038	1,038	1,038	1,038	1,038	1,038
Standard errors	robust	robust	cluster	cluster	cluster	cluster

Source: author's calculation using the NLSY79.
Notes: standard errors are clustered at the individual's level and are shown in parentheses. The signs *, **, *** indicates significance at 10, 5, and 1 percent. Other regression covariates not shown in the table are years of schooling, quartic function of experience, a dummy for marital status, a dummy for residence in SMSA, census region dummies, interaction terms between education and the census region dummies, and time dummies.

of schooling increases the prime-age earnings by 0.23 standard deviation.

I then analyze whether the marginal effects of the measures of school quality vary with the years of schooling. Following Betts (1995), I include interaction terms between the years of schooling and the measures of school quality:

$$\ln w_{it} = \alpha + \widetilde{\rho} S_{it} + \widetilde{\theta}\mathrm{Q}_i + \widetilde{\delta}\overline{S}_{it} \cdot \overline{\mathrm{Q}}_i + \gamma \mathrm{E}_{it} + \beta \mathrm{X}_{it} + \epsilon_{it}. \tag{1.3}$$

where $\overline{S}$ and $\overline{\mathbf{Q}}$ are schooling and school quality variables centered at their means, respectively. The centering of these variables allows us to interpret the estimated parameters conditional on the averages of the interacted variables. For instance, the estimated parameters $\tilde{\theta}$ capture the marginal effects of the measures of school quality for individuals with average years of schooling. This simple manipulation also allows us to compare these estimates with estimates using model Specification 1.2.

I present the marginal effects of the measures of school quality on earnings conditional on the average years of schooling in Column 4 and 6 Table 1.5. Note that the average years of schooling is about 12.8 which implies that on average the individuals are high-school graduates. There are no significant interaction effects between years of schooling and the measures of school quality and the inclusion of the interaction terms does not improve the fit of the model. Nevertheless, the result is similar to the one in the baseline model: the percentage of teachers with a Master's degree is significant only when individuals were older. This result is consistent with findings in Wachtel (1976) who used district-level per-student expenditure as a measure of school quality. He finds that the estimated parameter of the per-student expenditure is larger when individuals were older.

I re-estimate Specification 1.3 using different sets of controls to investigate whether the effects of the measures of school quality are consistent across specifications. For these estimations, I use the sample of individuals when they were 40 or older and I present the results in Table 1.6. First, the parents' level of education might be positively correlated with the measures of school quality. Individuals from better educated families may have lived in districts that host good schools and may have been motivated to choose better quality schools. Omitting parents' education from the specification might biased estimated parameters of the measures of school quality upward. Column 2 and 3 in Table 1.6 show that the estimated effect of a one percent increase in the percentage of teachers with a Master's degree drops from 0.26 percent to 0.20 percent after controlling for parents' level of education. Similarly, individuals of wealthier families might have lived in more affluent districts with better public schools. Omitting family wealth may also biased the measures of school quality upward. However, as shown in Column 4, the inclusion of the family wealth variable does not change the estimated parameter of the percentage of teachers with a Master's degree. This should not be surprising because family wealth and parents' level of education are strongly correlated.

Individuals' abilities might also bias the estimated parameters of the measures of school quality since the more able individuals might have enrolled in better-quality schools. Therefore, I use age-adjusted AFQT test score as a proxy of individuals' ability. The estimated parameter of the percentage of teachers with a Master's degree is not sensitive to addition of this variable. Lastly, the estimated parameters are not sensitive to addition of state fixed effects or state by year fixed effects in the model specification. These sensitivity analyses show that the effects of the percentage of teachers with a Master's degree are quite robust.

So far, I focus on teacher-student ratio, relative teacher salaries, and percentage of teachers with a Master's degree as the main measures of school quality. However, as shown in Section 1.3.2, these measures of school quality are correlated with other quality metrics such as the percentage of black students, the percentage of economically disadvantaged students, the percentage of teacher who left the school, the percentage of grade 10 dropouts, and library books per students. I investigate whether the estimated effects of the main measures of school quality are robust by augmenting Specification 1.2 with one quality metric at a time. I present the results of the estimations in Table 1.7. We can observe that the esti-

Table 1.6: The Estimated Effects of the Measures of School Quality on Earnings

	1: 40+ y.o.	2: 40+ y.o.	3: 40+ y.o.	4: 40+ y.o.	5: 40+ y.o.	6: 40+ y.o.	7: 40+ y.o.
Teacher-student ratio	2.111	2.309	2.477*	2.468*	2.255	1.313	1.360
	(1.318)	(1.438)	(1.457)	(1.464)	(1.438)	(1.478)	(1.495)
Relative teacher salaries	-0.076	-0.025	-0.055	-0.052	-0.063	-0.098	-0.060
	(0.138)	(0.143)	(0.145)	(0.145)	(0.145)	(0.164)	(0.164)
% of teachers with a Master's	0.00257***	0.00186**	0.00195**	0.00195**	0.00194**	0.00224**	0.00220**
	(0.000922)	(0.000942)	(0.000957)	(0.000963)	(0.000979)	(0.00104)	(0.00108)
Observations	3,146	3,032	2,942	2,933	2,862	2,862	2,862
Adjusted R^2	0.225	0.226	0.229	0.230	0.233	0.256	0.247
Regional Control	Region FE	Region FE	Region FE	Region FE	Region FE	State FE	StateXYear FE
Father's education		Y	Y	Y	Y	Y	Y
Mother's education			Y	Y	Y	Y	Y
Family's wealth in 1979				Y	Y	Y	Y
AFQT score					Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y
Standard Errors	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster

Source: author's calculation using the NLSY79.
Notes: standard errors are clustered at the individual level and are shown in parentheses. The signs *, **, *** indicates significance at 10, 5, and 1 percent. Other regression covariates not shown in the table are years of schooling, quartic function of experience, a dummy for marital status, a dummy for residence in SMSA, census region dummies, interaction terms between education and the census region dummies, time dummies, state dummies, and interaction terms between state and time dummies. The AFQT score is adjusted by the individuals' age.

mated coefficients of the percentage of teachers with a Master's degree are not sensitive to addition of other quality metrics. In all specifications, we reject the null that the percentage of teachers with a Master's degree does not affect individuals' earnings. It is interesting to note that the percentage of economically disadvantaged students has a negative effect on individuals' earnings and that the magnitude of its effect is similar to that of the percentage of teachers with a Master's degree.

1.4.2 The Effects of the Measures of School Quality Over the Life Cycle

The results so far suggest that the effect of the percentage of teachers with a Master's degree is higher when individuals were older. Thus, the logical next step is to estimate the effects of the measures of school quality over individuals' life cycle. I estimate a model with an interaction between schooling and age and an interaction between the percentage of teachers with a Master's degree and age:

$$\ln w_{it} = \alpha + \rho S_{it} + \theta \mathrm{Q}_i + \beta age_{it} + \theta S_{it} \cdot age_{it} + \delta \mathrm{Q}_i \cdot age_{it} + \gamma \mathrm{E}_{it} + \beta \mathrm{X}_{it} + \epsilon_{it}. \tag{1.4}$$

I report the results of the regressions in Table 1.8. I find a significant interaction effect between schooling and age and a significant interaction effect the percentage of teachers with a Master's degree and age. These estimates suggest that the effect of schooling and the effect of the percentage of teachers with a Master's degree increases as individuals became older. I find no interaction effect between age and the other measures of school quality. This suggests that the teacher-student ratio and the relative salaries of teachers effects on earnings are constant over individuals' life cycles.

It is important to obtain the effects of the measures of school quality at a certain age and to investigate whether the effects are significant or not. I proceed by calculating partial effects of schooling and partial effects of the percentage of teachers with a Master's degree

Table 1.7: The Estimated Effects of the Measures of School Quality on Earnings

	1: 17-32 y.o.	2: 40+ y.o.	3: 40+ y.o.	4: 40+ y.o.	5: 40+ y.o.	6: 40+ y.o.	7: 40+ y.o.	8: 40+ y.o.
Teacher-student ratio	0.176	0.906	0.767	0.829	0.852	0.841	0.841	0.664
	(1.020)	(1.201)	(1.193)	(1.211)	(1.191)	(1.193)	(1.193)	(1.205)
Relative teacher salaries	0.003	0.003	0.005	0.004	0.001	0.001	0.001	0.002
	(0.013)	(0.019)	(0.018)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
% of teachers with a Master's	0.000	0.002**	0.002**	0.002**	0.002*	0.002*	0.002*	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
% of black students			-0.003*	-0.003*	-0.003	-0.003	-0.003	-0.002
			(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
% of teachers who left				-0.001	-0.001	-0.001	-0.001	-0.001
				(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
% of disadvantaged students					-0.002	-0.002	-0.002	-0.002*
					(0.001)	(0.001)	(0.001)	(0.001)
% of grade 10 dropouts						0.000	0.000	0.000
						(0.001)	(0.001)	(0.001)
Library books per student								0.002
								(0.001)
Observations	10,250	3,640	3,640	3,640	3,640	3,640	3,640	3,640
Adjusted R-squared	0.364	0.217	0.219	0.219	0.220	0.220	0.220	0.220
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Standard Errors	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster

Source: author's calculation using the NLSY79.
Notes: the estimations use publicly available NLSY79 dataset. Standard errors are clustered at the individual level and are shown in parentheses. The signs *, **, *** indicates significance at 10, 5, and 1 percent. Other regression covariates not shown in the table are years of schooling, quartic function of experience, a dummy for marital status, a dummy for residence in SMSA, region dummies, interaction terms between education and the region dummies, and time dummies. The AFQT score is adjusted by the individuals' age.

at each age point:

$$\frac{\partial \ln w}{\partial S} = \rho + \theta \cdot \text{age} \qquad \frac{\partial \ln w}{\partial \text{master}} = \theta_{\text{master}} + \delta_{\text{master}} \cdot \text{age} \tag{1.5}$$

and the corresponding standard errors and 95 percent confidence intervals. I plot the partial effects of schooling and the percentage of teachers with a Master's degree and their corresponding confidence intervals in Figure 1.2. First, we can observe that the estimated partial effects increases monotonically with age. The percentage of teachers with a Master's degree has no effects on earnings when individuals were below 30. However, the percentage of teachers with a Master's degree has positive and significant effects on earnings when individuals were 30 or older. The pattern of the marginal effects of the percentage of teachers with a Master's degree is similar to the pattern of marginal effects of schooling. I also estimate the effects of the teacher-student ratio on earnings over the life cycle of individuals. The estimated marginal effects are positive when individuals were older however they are not statistically significant. The estimated marginal effects of relative teacher salaries on earnings are close to zero.[5]

The fact that the percentage of teachers with a Master's degree is significant for older individuals is an important finding for the literature. A direct comparison for the results in this study is the results in Betts (1996), who investigated the effects of the measures of school quality on earnings at ages 40-55. Betts (1996) uses earnings of individuals aged 40-55 in the 1980 Census to predict NLSY79 individuals' earnings when they are 40-55 years old. In his study, the estimated parameter of the percentage of teachers with a Master's degree is negative although it is statistically not different to zero.

I argue that there are two explanations for the difference. First, the 1980 Census is cross-sectional data, hence we can only observe one earning observation for each individual. One needs different individuals to construct a series of earnings data for age 40-55. How-

[5]See Figure A.1 for the marginal effects of teacher-student ratio and Figure A.2 for the marginal effects of relative teacher salaries.

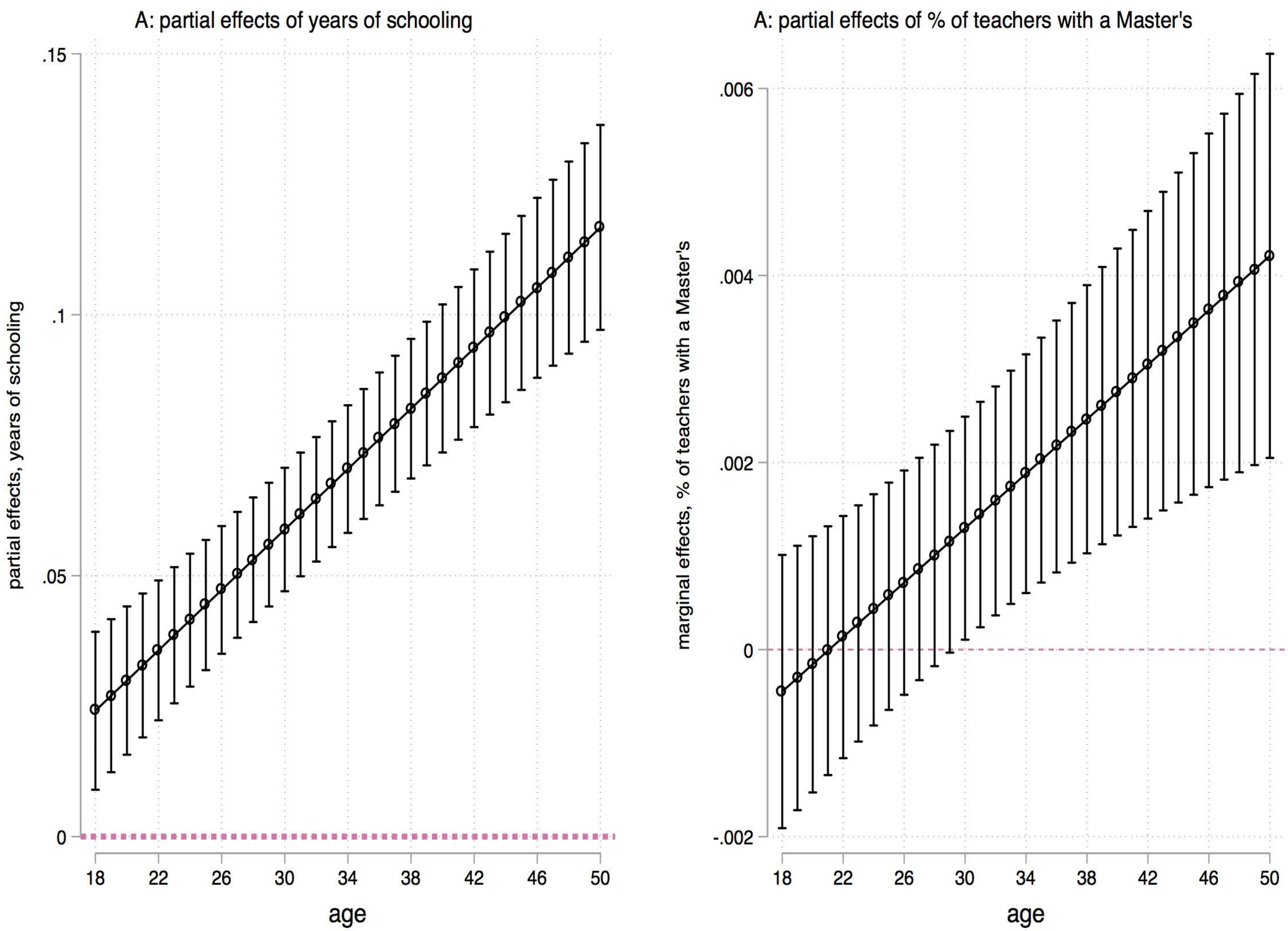

Figure 1.2: Partial Effects of the Years of Schooling and the Percentage of Teachers with a Master's Degree by Age. Source: author's calculation using the NLYS79. Notes: standard errors are clustered at the individual's level. The covariates used in the estimation model are years of schooling, quartic function of experience, a dummy for marital status, a dummy for residence in SMSA, census region dummies, interaction terms between education and the census region dummies, and time dummies.

Table 1.8: The Effects of Schooling and the Measures of School Quality Over the Life Cycle

Measures of school quality	1: Model 1	2: Model 2	3: Model 3
Education	-0.028**	-0.080***	-0.107***
	(0.013)	(0.014)	(0.021)
Teacher-student ratio	-1.272	-1.276	-1.336
	(2.048)	(2.019)	(2.062)
Relative teacher salaries	0.007	-0.003	-0.001
	(0.031)	(0.031)	(0.031)
% of teachers with a Master's	-0.003**	-0.003**	-0.003**
	(0.001)	(0.001)	(0.001)
Age	0.072***	0.029***	0.034***
	(0.001)	(0.003)	(0.007)
Education · age	0.003***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)
Teacher-student ratio · age	0.065	0.064	0.059
	(0.060)	(0.063)	(0.063)
Relative teacher salaries · age	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
% of teachers with a Master's · age	0.0001***	0.0001***	0.0001***
	(0.000)	(0.000)	(0.000)
Observations	17,240	16,366	16,366
Controls	N	Y	Y
Region and time dummies	N	N	Y

Source: author's calculation using the NLSY79.
Notes: these estimations are produced using publicly observed dataset. standard errors are clustered at the individual's level and are shown in parentheses. The signs *, **, *** indicates significance at 10, 5, and 1 percent. Other regression covariates not shown in the table are years of schooling, quartic function of experience, a dummy for marital status, a dummy for residence in SMSA, region and time dummies.

ever, each individual has different observed as well as unobserved characteristics, which may significantly affect earnings in different ways. Second, the earnings obtained from the 1980 Census cannot capture the actual state of the economy and technological progress, which significantly affect occupation-specific earnings and earnings growth. Therefore, earnings data for a specific occupation in the 1980 Census are not comparable to the post-1990 earnings data in the NLSY79.

Table 1.9: The Effects of the Measures of School Quality by Highest Years of Schooling Completed

Sample: 17-32 years	1: HS dropout	2: HS graduate	3: Some College	4: 4Y College	5: Post-Graduate
Teacher-student ratio	0.512	0.0845	2.440	1.109	-2.381
	(1.391)	(1.507)	(2.670)	(2.379)	(2.802)
Relative teacher salaries	0.0278	0.0212	0.00532	-0.00805	-0.0267
	(0.0183)	(0.0207)	(0.0299)	(0.0299)	(0.0302)
% of teachers with a Master's	0.00108	0.00123	-0.000221	0.00115	-0.00128
	(0.000766)	(0.000852)	(0.00151)	(0.00159)	(0.00202)
Observations	5236	4595	2022	1669	1323
Controls	Y	Y	Y	Y	Y
Standard Errors	Cluster	Cluster	Cluster	Cluster	Cluster
Sample: 40+ years	1: HS dropout	2: HS graduate	3: Some College	4: 4Y College	5: Post-Graduate
Teacher-student ratio	1.312	0.816	0.192	-1.382	2.935
	(1.520)	(1.607)	(2.713)	(3.254)	(4.382)
Relative teacher's salary	0.0351	0.0173	0.0224	-0.00116	-0.0609*
	(0.0277)	(0.0309)	(0.0432)	(0.0282)	(0.0320)
% of teachers with a Master's	0.00246**	0.00279**	0.00143	0.000519	0.00158
	(0.00108)	(0.00117)	(0.00178)	(0.00254)	(0.00264)
Observations	1828	1623	744	565	503
Controls	Y	Y	Y	Y	Y
Standard Errors	Cluster	Cluster	Cluster	Cluster	Cluster

Source: author's calculation using the NLSY79.
Notes: standard errors are clustered at the individual level and are shown in parentheses. The signs *, **, *** indicates significance at 10, 5, and 1 percent. Other regression covariates not shown in the table are years of schooling, quartic function of experience, a dummy for marital status, a dummy for residence in SMSA, census region dummies, interaction terms between education and the census region dummies, and time dummies.

1.4.3 The Effects of the Measures of School Quality by Years of Schooling Completed

The results so far show that the percentage of teachers with a Master's degree indeed affect earnings and the effects are stronger as individuals became older. In this section, I analyze whether the effects of the measures of school quality differ by years of schooling completed. I identify the highest years of schooling that each individual completed. I then estimate the model in Specification 1.2 by highest years of schooling completed: less than high school, high school, some college, four-year college, and post-graduate. For each schooling group, I estimate the model using the sample when individuals were 17-32 years old and the sample when individuals were 40 years or older. Thus, for each schooling group, I have the same group of individuals in both samples.

I report the results in Table 1.9. Consistent with the previous results, the measures of school quality have no effects on earnings when individuals were young irrespective of the schooling group. On the other hand, the percentage of teachers with a Master's degree affect earnings when individuals were older but only among individuals who did not finish high school and individuals who completed high-school. This is an interesting finding because individuals with at most a high school diploma attended high schools with relatively lower share of teachers with a Master's degree. The average percentage of teachers with a Master's degree for individuals with at most a high school degree is 42.8 percent while the average for individuals with a 4-year college degree is 49.6 percent. This result suggests that attending a better-quality school matters for individuals who did not attend college.

1.4.4 Does Return to Schooling Vary by Measures of School Quality?

I now investigate whether the marginal effects of the the return to schooling vary with the measures of school quality. This analysis is important as the state-level studies in the literature. For example, Card and Krueger (1992a) find that the measures of school quality have significant relationships with the return to education. For this analysis, I categorize each measure of school quality into quartiles and estimate Specification 1.2 for each quartile. I estimate the model using the sample when individuals were 40 years or older. Table 1.10 presents the estimates of the return to schooling for each quartile of the measures of school quality. The results show that individuals who attended schools with the lowest teacher-student ratio or individuals who attended schools with the highest percentage of teachers with a Master's degree earned higher returns to schooling. One of the explanations is that the percentage of teachers with a Master's degree is negatively correlated with the teacher-student ratio.

Table 1.10: Return to Schooling by the Measures of School Quality

	1: 1st Quartile	2: 2nd Quartile	3: 3rd Quartile	4: 4th Quartile
Return to schooling x Teacher-student ratio	6.406** (3.233)	9.614 (13.14)	4.288 (6.927)	0.434 (0.908)
Observations	749	778	814	805
Return to schooling x Relative teacher salaries	0.182 (0.117)	0.0155 (0.162)	-0.0349 (0.218)	0.0219 (0.122)
Observations	792	832	772	750
Return to schooling x % of teachers with a Master's	0.158 (0.183)	0.0251 (0.182)	0.198 (0.287)	0.453** (0.194)
Observations	783	810	781	772
Controls	Y	Y	Y	Y
Standard Errors	Cluster	Cluster	Cluster	Cluster

Source: author's calculation using the NLSY79.
Notes: standard errors are clustered at the individual level and are shown in parentheses. The signs *, **, *** indicates significance at 10, 5, and 1 percent. Other regression covariates not shown in the table are years of schooling, quartic function of experience, a dummy for marital status, a dummy for residence in SMSA, census region dummies, interaction terms between education and the census region dummies, and time dummies.

1.5. CONCLUSION

There is an open debate in the literature that looks at the effects of the measures of school quality on labor market earnings. Studies using state or aggregate-level measures of school quality such as teacher-student ratio and relative teacher wage find positive effects on labor market earnings. On the other hand, studies using school-level measures of school quality, such as Betts (1995), find no effect of the measures of school quality on labor market earnings. Card and Krueger (1996) argue that Betts' results are driven by the use of a relatively young sample. A follow-up paper by Betts (1996) addresses this criticism by using the 1980 Census to predict prime-age earnings of individuals in the NLSY79 sample. He finds no significant effects of the measures of school quality on earnings. While this study supports Betts (1995), I argue that there are many drawbacks from the wage-matching method used in this study.

This study reopens this debate by analyzing the direct effects of the measures of school quality on earnings when individuals were older. The more recent NLSY79 dataset provides an opportunity to analyze the effects of the measures of school quality when individuals were 40 or older. Specifically, the dataset allows comparison of the effects of the measures of school quality when individuals were young and when they were older. This analysis is essential to address Card and Krueger's claim that the use of a relatively young sample in Betts (1995) understates the effects of the measures of school quality.

I find strong evidence that the measures of school quality affect earnings only when individuals are 40 or older. The estimates are quite consistent across model specifications even after controlling family characteristics and a measure of individuals' ability. The point estimates suggest that a one standard-deviation increase in the percentage of teachers with a Master's degree increases the prime-age earnings of those with the average years of schooling by 0.06 to 0.07 standard deviation. I also find that the estimated effects of the percentage of teachers with a Master's degree is decreasing with the years of schooling, and it is significant for individuals who completed high school or one-year college.

The results in this study have strong implications on the literature. First, the results support Card and Krueger's claim that the measures of school quality affect earnings only after individuals are older. Second, the results in this study reconcile the findings between studies that used state- and school-level data. Together with state-level studies, the results in this study conclude that the measures of school quality affect individuals' earnings in labor market.

2. MORE THAN DOLLARS FOR SCHOLARS: THE IMPACTS OF THE DELL SCHOLARS PROGRAM ON COLLEGE ACCESS, PERSISTENCE, AND DEGREE ATTAINMENT

Co-authored with Lindsay C. Page, Benjamin L. Castle, and Stacy S. Kehoe

2.1. INTRODUCTION

Over the last several decades, a variety of organizations—from local college access programs to the federal government—have invested hundreds of billions of dollars in programs and policies to improve college outcomes for economically-disadvantaged youth. College enrollment rates have increased substantially over this period. Yet, socioeconomic gaps in college completion have actually widened. For example, while the share of young people in the top income quartile earning bachelor's degrees by age 25 increased from 36 to 54 percent between the 1961-1964 and 1979-1982 birth cohorts, degree attainment by age 25 among students in the lowest income quartile only increased from five to nine percent over the same time period (Bailey and Dynarski, 2011).

An extensive body of rigorous empirical research demonstrates that a variety of college access efforts can generate substantial improvements in college entry for lower-income populations.[1] For example, researchers have found positive enrollment effects from need-based grant programs administered by state and federal governments (Castleman and Long,

[1] Page and Scott-Clayton (2016) provides a comprehensive review of the literature.

2016; Dynarski, 2003; Kane, 2003); merit scholarships that reward academic achievement in high school (Dynarski, 2008; Scott-Clayton, 2011); college advising programs that provide students with individualized support to complete college and financial aid applications (Avery, 2013; Carrell and Sacerdote, 2013; Castleman and Goodman, 2016); and informational campaigns that provide students with simplified information about college and financial aid, reminders to complete important tasks, and access to advising when students need assistance (Castleman and Page, 2014b,a, 2015; Castleman et al., 2014; Hoxby et al., 2013).

A series of recent studies have also demonstrated positive effects of financial, advising, and informational interventions on college persistence and completion. Some of these interventions target students before they enter higher education. Bettinger et al. (2012) show that providing low-income families with assistance completing the Free Application for Federal Student Aid (FAFSA) during the income tax preparation process leads to substantial increases in the share of students that enroll and persist for at least two years in college. In addition, financial aid efforts directed to students based on financial need (Castleman and Long, 2016) and and on academic merit (e.g. Scott-Clayton, 2011) have improved rates of bachelor's degree attainment. Not all scholarship programs, however, demonstrate positive long-term effects for students. For example, DesJardins and McCall (2014) find that while the Gates Millenium Scholars program, which is awarded to high-achieving, low-income students, led to modest increases in students' GPA through junior year of college, it had no impact on bachelor's degree attainment.

Other programs provide outreach and support to students after they have begun college. Castleman and Page (2016) provide experimental evidence that targeted text-based reminders about re-application for financial aid can improve first-to-second year persistence. Other persistence interventions are more comprehensive. Inside Track, for example, is a private company that contracts with colleges to provide students with coaching (primarily delivered via phone) about issues and challenges that arise over the course of the academic year. Freshmen who were randomly assigned to receive a one-on-one, sustained college coaching from Inside Track were 4 percentage points more likely to earn a degree than students who did not receive this coaching—a 13 percent relative increase (Bettinger and Baker, 2014). More comprehensive still is the City University of New York's Accelerated Study in Associates Program (ASAP), which provides intensive structural advising and financial support for selected community college students. Those who were randomly assigned to participate in ASAP were 66 percent more likely to earn a degree within several years than their peers who were not selected to participate (Scrivener and Weiss, 2013). Angrist et al. (2014) examine the impact of scholarship support from the Susan Thompson Buffett Foundation. Students selected as Buffett Scholars receive generous financial aid and, in some cases, academic and social supports through on-campus learning communities at selected colleges in Nebraska. Experimental evidence reveals that the program has sizable impacts on both institutional choice and early college persistence. Recipients are substantially more likely to matriculate in a four-year college and are more likely to persist into their sophomore year of college. Finally, Clotfelter et al. (2016) find that the Carolina Covenant, which provides low-income students admitted to the University of North Carolina at Chapel Hill with a full cost of attendance scholarship and additional counseling and supports, led to improvements in on-time bachelor's degree attainment on the order of 8 percentage points.

We contribute to the growing but still nascent literature on the long-term effects of interventions focused on college completion by investigating the impact of the Michael and Susan Dell Foundation's Dell Scholars Program. The Dell Scholars Program provides a combination of financial support and individualized advising to scholarship recipients, both as they enter college and throughout the duration of their postsecondary enrollment. Like the Buffett program, the Dell approach is resource intensive; Dell Scholars are awarded a one-time scholarship support of up to $20,000 in addition to the operating costs of the program itself. This programmatic design is motivated by a theory of action that, in order to meaningfully increase the share of lower-income students who earn a college degree, it is necessary both to address financial constraints students face and to provide ongoing support for the academic, cultural and other challenges that students experience during their college careers. Indeed, in a recent review on factors predicting college completion, Perna and Jones (2013) identify three key aspects of the postsecondary experience: college financing, academic achievement and social integration. Thus, providing a combination of financial supports and other wrap-around supports that target the non-financial domains of student success may be a more effective and efficient approach than offering either in isolation Page and Scott-Clayton (2016).

We isolate the unique impact of the Dell Scholars Program on college completion by capitalizing on an arbitrary cutoff in the selection process that determines which applicants are selected as Dell Scholars. Using a regression discontinuity (RD) design, we find that while being selected a Dell Scholar has no impact on initial college enrollment or early college persistence, scholars at the margin of eligibility appear more likely to persist into the third year of college and are significantly more likely to earn a bachelor's degree on-time or six years after high school graduation. Specifically, students just above the margin of scholar selection are 6-9 percentage points more likely to earn a bachelor's degree within four years and 16 percentage points more likely to earn a bachelor's degree within six years compared to their counterparts who just missed being named a Dell Scholar. These impacts are sizable and represent a nearly 22 percent or greater increase in both four- and six-year bachelor's attainment.

In addition to our impact analyses, we conduct a back-of-the-envelope cost-benefit analysis to assess whether these substantial increases in college completion are sufficient to merit the intensive investment that the Dell Scholars Program makes in its recipients. Although our calculations hinge on several assumptions, as we outline, they nevertheless suggest that the investment in Dell Scholars has a positive rate of return. Given that those selected as Dell Scholars are predominantly first-generation college students from low-income backgrounds, our findings have important implications for efforts to expand college success in the US. We structure the paper as follows. In Section 2.2 we describe the Dell Scholar's program, including their application and selection procedure. In Section 2.3 and Section 2.4, we highlight the data and analytic strategies we bring to bear in our investigation. In Section 2.5 we present our results, and in Section 2.6 we conclude.

2.2. DELL SCHOLARS PROGRAM

2.2.1 Description

The Dell Scholars Program is a unique college success initiative sponsored and administrated by the Michael and Susan Dell Foundation. The program targets motivated low-income students who have the potential to enroll and succeed in college. Students selected to be Dell Scholars receive generous financial support towards the costs of higher education. This includes a total of up to $20,000 in scholarship funds, a laptop computer and textbook support. We provide a back-of-the envelope calculation to show that the funds can save scholars substantial working time. We make a simplifying assumption that a scholar completes a bachelor degree in four years and that a scholar would earn $10 an hour from working part time as a college student. This implies that a scholar is able to forgo about 500 working hours every year. This figure is quite substantial because students between the age of 20 and 25 in the 1979 National Longitudinal Survey of Youth data on average worked 700 hours a year.

Compared to other scholarship programs, the Dell effort is relatively unique in that it also provides ongoing outreach, close monitoring, and assistance to scholars, even though they are geographically dispersed to postsecondary institutions across the US. As the program materials explain, beyond formal scholarship funding, the program also provides: "*... an ongoing support and assistance to address all of the emotional, lifestyle, and financial challenges that may prevent our scholars from completing college. These pressures range from dealing with stress, to getting out of debt, to managing child care, and dealing with life circumstances as they arise.*"[2]

[2]Text provided by Oscar Sweeten Lopez, Portfolio Director for the Dell Scholars program, January 27,

This ongoing support is actualized by requiring scholars to input postsecondary progress information into a sophisticated data and communication portal that is closely monitored by program staff. Scholars input information about key college success metrics including academic performance, financial aid, and social integration.[3] Scholars are required to report information via the portal prior to postsecondary enrollment, at the end of fall and spring semesters freshman year, and once annually after the first year of college. The portal is designed to flag inputs associated with threats to college persistence and to immediately trigger the process of providing individualized follow-up, support, and guidance. This data-driven program model allows a small program staff to provide proactive, intensive social support to scholars who are at risk of attrition at any point during their postsecondary trajectory.

Since 2004, the program has selected and supported over 3,000 scholars. Currently, it selects approximately 300 students as Dell scholars annually. Despite the small annual cohort size, the Dell Scholars program is well known. Between 2009 and 2012, for example, the program selected a total of 1,201 scholars from a pool of 23,600 applicants.

2.2.2 Scholar Application

Students apply to be Dell Scholars during their senior year of high school. To be eligible, students must meet certain preliminary criteria. First, students must have participated in one of several college readiness programs during the last two years of high school.[4] In addition, applicants must be graduating from an accredited high school, earn a minimum 2.4 grade point average (GPA), be financially eligible to receive a federal Pell grant in the first year of college, and plan to enroll full time in a four-year college at an accredited higher

2016.

[3] In addition to this student self-reported information, students are required to submit documentation, such as transcripts, which are used to verify the student-reported data.

[4] At the time of our writing, these programs included Alliance College-Ready Public Schools, AP Strategies, Aspire, AVID, Bottom Line, Breakthrough Austin, College Forward, Cristo Rey Network, Fulfillment Fund, GEAR UP, Genesys Works, Green Dot Public Schools, IDEA Academy, KIPP Academy/ KIPP Through College(KTC), Mastery Charter Schools, Noble Charter, One Goal, Philadelphia Futures, Upward Bound, Upward Bound - Math Science, YES Prep Public Schools, Uncommon Schools, and Uplift Education.

education institution in the fall directly after high school graduation.

Qualified students complete an online application that gathers information about high school grades, test scores, the college readiness program (CRP) in which the student participated, college plans, home and work responsibilities, financial information, and home environment. The online application form can be found on Dell Scholars Program website.[5] Each year, eligible high school seniors can apply to the Dell Scholarship Program between November 1 and January 15. After this date, the application is closed, and the selection process begins.

2.2.3 Selection Process and Selection Algorithm

Scholar selection proceeds in two phases. The first phase is the selection of semifinalists from among all qualified applicants, and the second phase is the selection of scholars from among semifinalists. The Dell Scholars program assesses prospective scholars based on three main criteria referred to as GPA: Grit, Potential, and Ambition. In each phase, the program scores students numerically along three dimensions: academics, disadvantage, and responsibility. These dimensions along with the eligibility criteria map directly onto the Grit-Potential-Ambition framework. Participating in a college readiness program and having a plan to enroll in a four-year college show an applicant's ambition. The academic dimension, which assesses academic achievements in high school, measures the applicant's potential. The final criterion, grit, is intended to target students who have overcome personal challenges in their lives related to their families, schools or communities. This criterion is assessed with the measures of disadvantage and student responsibility. Each of these dimensions includes several sub-categories. For instance, the academic dimension consists of an academic difficulty index, course count, and high school grade point average while the disadvantage dimension consists of parents' income, parents' level of education, enrollment in state or federal aid programs, and student living situation.

[5] See https://apply.dellscholars.org/Application/Print.

The Dell Scholars Program utilizes scoring algorithms, one for each selection phase, to compute overall scores. We refer to these as the semifinalist algorithm and the scholar algorithm, respectively. The program uses the semifinalist algorithm to compute a final application score for each student who starts an application. The semifinalist algorithm consists of three sub-algorithms: a sub-category scoring algorithm, a calibration algorithm, and a final score algorithm. The first algorithm identifies responses for each question in a particular sub-category and computes the score for the sub-category. The calibration algorithm standardized the sub-category scores and computes adjusted weights for each sub-category. The final score algorithm use the adjusted weights and the standardized sub-category scores to calculate the category score. The program then uses the category scores and repeats the algorithms to compute the final application score.

Students are then ranked on this final application score, and the top 900 students are selected as semifinalists. Semifinalists are notified on February 1 and are then required to provide additional application materials, including a high school transcript, a Student Aid Report obtained after completing the Free Application for Federal Student Aid (FAFSA), responses to additional short-answer questions, and a letter of recommendation before March 10. The semifinalists who complete these requirements are referred to as finalists and enter the scholar selection process. Finalist applications are distributed among and reviewed by a selection committee consisting of approximately 60 members. Each finalist's full application is reviewed and scored by two readers. The assignment process ensures that both readers in the pair have zip codes different from the finalists they are reviewing. Each reader in the pair individually reviews each assigned complete application, including recommendation letters, and scores each item in the application.

At the end of March, the readers submit all application reviews. Super-readers, a subset of readers with extensive experience in scoring applications, review and score applications that need an additional evaluation because first two readers awarded scores that deviated substantially from each other. Once all applications have been reviewed and scored, the

Table 2.1: Categories and Corresponding Weights in the Scoring Algorithms

Category	**Semifinalist Algorithm**	**Finalist Algorithm**
Academics	0.28	0.34
Disadvantage index	0.28	0.34
Responsibility: home	0.18	0.16
Responsibility: work	0.18	0.16
Responsibility: community	0.08	

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.

program uses the scholar algorithm to compute a final score for each finalist. The scholar algorithm also consists of three sub-algorithms: a sub-category scoring algorithm, a calibration algorithm, and a final score algorithm. However, there are notable differences between the semifinalist and finalist algorithms. First, the categories and the corresponding weights are different, as summarized in Table 2.1. Second, the calibration algorithm in the scholar selection process computes score adjustments before it standardizes each score. The calibration routine in the finalist algorithm has this extra step because two selection committee members review each application. Specifically, the algorithm adjusts a committee's score for a sub-category if the score difference with her partner lies outside of a specific computed interval. The last difference is the computation of the final score. In the finalist algorithm, a category score is included in the computation of the final score if the score lies within a program-specified interval. Thus, the final score can include only one category score if the other scores lie outside of the interval.

The program ranks the finalists based on the final scores, and the top 300 finalists are selected as Dell Scholars. The program announces the selected Dell Scholars on April 10 of each year.

2.2.4 Research Questions

Our main objective is to evaluate the impact of the Dell Scholars Program on immediate college enrollment, persistence and completion. In addition, we investigate how these impacts vary by observable student characteristics, such as status as a would-be first-generation college student. Our investigation is guided by the following research questions:

1. What is the impact of being selected as a Dell Scholar on college enrollment, persistence and degree completion?

2. To what extent do these impacts vary by salient student-level characteristics?

3. Do the benefits of the Dell Scholars program justify the costs?

2.3. DATA

The selection process implies that there is a cohort-specific cutoff score that determines whether a finalist was admitted as a Dell Scholar. Specifically, finalists whose score were higher than the 301st ranked scholar were admitted as scholars. This process lends itself perfectly to a regression discontinuity design for assessing the program's impact. In this section, we lay out the data and the analytic strategy for informing our research questions.

2.3.1 Data

Our data come from the Dell Scholar applicant records and the National Student Clearinghouse (NSC). The Dell Scholar applicant records provide comprehensive information about each applicant from the high school graduating classes of 2009 through 2012. The application data provide basic demographic information, such as gender, race and ethnicity, state of residence, and parents' education level and employment status. The data also include indicators of students' academic background, including standardized test scores, high school achievement, participation in college readiness programs, and top three postsecondary institution preferences (as students apply to be Dell Scholars prior to being admitted to any institutions to which they have applied).

While optional and not factored into the scholar selection process, the large majority of applicants took and reported scores for either the SAT or ACT test. We convert SAT test scores (critical reading, math, and writing) to ACT composite test scores using the ACT-SAT concordance table retrieved from the ACT website.[6] Information on applicants' high school achievement includes overall high school GPA as well as information on courses taken and course-level grades earned. Applicants also provide information about their responsibilities at home, work, and in their community. Lastly, the application data provide measures of applicants' financial circumstances, including household income and enrollment in state or federal aid programs.

In Table 2.2, we provide detailed counts of applicants across the 2009 through 2012 cohorts. Across these years, the program experienced a substantial growth in applications, with the 2012 applicant cohort being 39 percent larger than that of 2009. Across all years, selected semifinalists complete the finalist application process with a high rate of compliance. On average, nearly 90 percent of the semifinalists submitted the required documents and achieved finalist status for the selection of scholars each year. The final number of selected scholars varies minimally from the target of 300 annually. Note that we are missing 3 applicants from the 2010 cohort whose final application scores were missing. We are also missing 139 finalists from the 2011 cohort, one of whom is a scholar, due to an unknown system issue. This explains the discrepancy in the number of selected scholars in 2011.

Table 2.2: Counts of Applicants Across Cohorts with Non-missing Algorithm Scores

Cohort	**Total Applicants**	**Semifinalists**	**Finalists**	**Scholars**
2009	4,912	775	643	300
2010	5,340	921	811	301
2011	6,533	900	760	299
2012	6,815	901	805	301
Total	23,600	3,497	3,019	1,201

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.

To examine students' college-going outcomes, we use on college enrollment and college

[6]See http://www.act.org/aap/concordance/pdf/reference.pdf.

outcome data from the National Student Clearinghouse (NSC), a non-profit organization that maintains postsecondary enrollment records for approximately 96 percent of colleges and universities in the US.[7] The NSC data provide semester-level enrollment information at the student level. These data allow us to observe whether and where students were enrolled in college. Additionally, the data provide classifications related to students' postsecondary institutions such as whether they are public or private institutions and whether they are two or four-year colleges. Finally, we are able to observe whether students progressed through to degree completion. Taken together, data from the NSC provides a comprehensive set of outcomes related to college enrollment, persistence and degree attainment.

In the NSC data, some applicants have overlapping enrollment records in a particular semester. For instance, an applicant could be enrolled in both a 2 and 4-year institutions in a particular semester. We implement decisions rules to clean the dataset from overlapping enrollment records in the following order.

1. We drop any enrollment record if an applicant withdrew from college, has a leave of absence status, or was deceased.

2. If a record indicates enrollment in a four-year college and the other record indicates enrollment in a two-year college, we select the former record.

3. There are 8 enrollment status: full-time, three-quarter-time, half-time, less-than-half-time enrollment. We select an enrollment record with a higher enrollment intensity.

4. If the enrollment records are equivalent in terms of intensity and the type of institution, we select the record in which the institution matches with the institution in the previous enrollment records.

5. All else equal, we select an enrollment record at random.

[7]The National Student Clearinghouse represents the best, comprehensive source of college enrollment information for US students. Nevertheless, the overall coverage is imperfect and that coverage rates vary across states (Dynarski et al., 2015).

2.3.2 Construction of Outcome Variables

We use the cleaned NSC data to construct various college outcomes such as an immediate 4-year college enrollment, second-year persistence rate, third-year persistence rate, earning a 4-year bachelor degree on time, and earning a 4-year bachelor degree in 6 years. We are interested in the immediate enrollment in a four-year college because Dell Scholars Program selects applicants who planned to enroll in a bachelor's degree program in the fall directly after high school graduation.

Table 2.3: Matching Rates Between the Dell Applicants Records and the NSC Data

Cohort	**Dell Data**		**NSC Data**		**Percentage**	
	Non	**Semifinalist**	**Non**	**Semifinalist**	**Non**	**Semifinalist**
2009	4,137	775	3,902	748	94.32	96.52
2010	4,419	921	4,084	865	92.42	93.92
2011	5,633	900	5,204	834	92.38	92.67
2012	5,914	901	5,347	828	90.41	91.90
Cohort	**Dell Data**		**NSC Data**		**Percentage**	
	Non	**Scholar**	**Non**	**Scholar**	**Non**	**Scholar**
2009	343	300	332	290	96.79	96.67
2010	510	301	475	289	93.14	96.01
2011	461	299	418	276	90.67	92.31
2012	504	301	469	279	93.06	92.69

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.

We could not match several Dell Scholars Program applicants to the NSC data. We assume that these applicants did not attend college. Thus, we assign a value of zero to these applicants' outcome observations. In Table 2.3, we find that matching rates between the Dell and the NSC data are quite high. More than 90 percent of the Dell Scholars Program applicants and finalists are observed in the NSC data. The matching rates are quite close to average coverage rates reported in Dynarski et al. (2015). We also find that matching rates are quite balanced across statuses.

2.3.2.1 Immediate 4-Year College Enrollment Rate We apply the following decision rules to generate immediate college enrollment rate:

1. For each applicant in cohort t, we keep records of enrollment in the fall of year t.

2. We keep enrollment records if applicants were enrolled in a 4-year college.

3. We assign a value of 1 if applicants were immediately enrolled in a 4-year college after high-school graduation and a value of 0 otherwise.

4. We compute the average value to obtain the immediate college enrollment rate.

2.3.2.2 Persistence Rate We apply the following decision rules to generate second-year persistence rate:

1. For each applicant in cohort t, we keep records of enrollment in the fall and spring of year t and in the fall of year $t + 1$.

2. We keep enrollment records if applicants were enrolled in a 4-year college.

3. We keep applicants who kept at least three enrollment records between the fall of year t and the fall of year $t + 1$. Note that applicants who were enrolled in colleges with a quarter system may have more than three enrollment records between the fall of year t and the fall of year $t + 1$.

4. We assign a value of 1 if applicants persisted into the second year in a 4-year college and a value of 0 otherwise.

5. We compute the average value to obtain the second-year persistence rate.

We use the same logic to generate the third persistence rate. We keep records of enrollment in the fall- and spring-semester of year t and $t + 1$, and records of enrollment in the fall of year $t + 2$. We keep applicants who kept at least five enrollment records between the fall of year t and the fall of year $t + 2$. We assign a value of 1 if applicants persisted into the third year in a 4-year college and a value of 0 otherwise.

2.3.2.3 Bachelor Degree Completion Rate We apply the following decision rules to generate on-time bachelor degree completion rate:

1. We keep applicants who graduated from a 4-year college with a bachelor degree. The list of bachelor degrees that applicants earned include, but not limited to, Bachelor of Arts (BA or AB), Bachelor of Science (BS), Bachelor of Science in Nursing (BSN), Bachelor of Interdisciplinary Study (BIS), Bachelor of Business Administration (BBA), Bachelor of Fine Arts (BFA), Bachelor of Music (BM), Bachelor of Architecture, Bachelor of Education, and Bachelor of Social Work (BSW).

2. We assign a value of 1 if applicants earned a bachelor degree within 48 months and a value of 0 otherwise.

3. We compute the average value to obtain the bachelor degree completion rate.

We use a similar procedure to generate 6-year bachelor degree completion rate. We assign a value of 1 to applicants who earned a bachelor degree within 72 months and a value 0 otherwise.

2.3.3 Sample Characteristics

In Table 2.4, we present descriptive statistics of applicants' demographic characteristics, test score performance, and financial aid eligibility, pooled across cohorts. In the first column of Table 2.4, we present results overall, and in the remaining columns, we present results disaggregated by applicant status (e.g., non-semifinalist, semifinalist, finalist, and scholar). While not reported in the table, rates of missingness on student demographics are very low, ranging from 0 to 4 percent across most items. Missingness was most prevalent for SAT/ACT scores (nearly 16 percent), presumably for those students who either did not take a college entrance test or simply opted not to report their scores in their application.[8] About 3 percent

[8] Missingness for SAT/ACT scores is quite low among the finalists.

of applicants are missing their parents' education, with students less likely to report father's than mother's education. This may be an indication of students living in a single-parent (e.g., mother-headed) household.

Of applicants, 70 percent are female, three-quarters are either black or Hispanic, and nearly sixty percent are would-be first-generation college goers. Applicants exhibit an average ACT composite score of 20.15, which corresponds to approximately the 48th percentile of performance among all test takers. Those ultimately selected as scholars are similar in terms of gender and race but are even more likely to be first-generation college goers and have an average ACT performance of about 22, corresponding to approximately the 62nd percentile of the national distribution. Scholars also achieved a slightly higher high school GPA, on average. Therefore, the scholar selection process favors those applicants who are higher performing but from lesser means. As an additional indicator of this final point, while 75 percent of all applicants qualified for subsidized school meals, nearly all of eventual scholars did so.

In Table 2.5, we list the set of college-going outcomes, such as immediate college enrollment, college persistence, and college completion, the cohorts for which we examine these outcomes, and the average values of these outcomes, disaggregated by applicant's status. We can observe that the proportion of scholars who persisted into the third year is 76 percent while the proportion of non-scholars who persisted into the third year is only 61 percent. The proportion of scholars who completed a bachelor degree on time is 34 percent higher than the proportion of non-scholars who do so. Overall, this descriptive evidence reveals a consistent pattern of better college-going outcomes among Dell Scholars compared to non-scholar finalists as well as to the applicant pool overall. While being selected as a Dell Scholar may be the driver of these differences, they may also be attributable to differences in the characteristics of students ultimately selected as scholars, such as their higher levels of prior academic achievement. Therefore, we turn to discussing our analytic strategy for disentangling these possibilities in the next section.

Table 2.4: Summary Statistics of All Applicants, Overall and by Applicant Status, 2009-2012

Variables	**All**	**Non-Semifinalist**	**Semi-finalist**	**Finalist**	**Non-Scholar**	**Scholar**
Age	18.26	18.25	18.32	18.32	18.31	18.32
	(0.49)	(0.48)	(0.53)	(0.53)	(0.52)	(0.54)
Scaled GPA	0.85	0.84	0.89	0.90	0.88	0.92
	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
ACT equivalent	20.15	19.93	21.38	21.49	20.92	22.33
	(3.84)	(3.78)	(3.92)	(3.90)	(3.76)	(3.95)
Female	0.70	0.69	0.71	0.71	0.71	0.71
Asian	0.10	0.09	0.15	0.16	0.15	0.17
Black	0.23	0.23	0.24	0.23	0.23	0.23
Caucasian	0.16	0.16	0.13	0.12	0.12	0.14
Hispanic	0.53	0.53	0.47	0.48	0.50	0.45
Other Ethnicity	0.04	0.04	0.04	0.04	0.04	0.04
Received lunch program	0.75	0.72	0.95	0.96	0.96	0.96
Received food stamp	0.24	0.19	0.53	0.53	0.54	0.53
Enrolled in WIC	0.09	0.08	0.17	0.18	0.19	0.16
Enrolled in TANF	0.03	0.02	0.08	0.09	0.08	0.09
Enrolled in LIHEAP	0.10	0.08	0.21	0.21	0.20	0.23
Enrolled in SSI	0.08	0.06	0.17	0.18	0.18	0.18
Enrolled in free housing	0.06	0.05	0.16	0.16	0.16	0.17
Enrolled in SSD	0.10	0.09	0.17	0.17	0.15	0.19
Enrolled in health insurance	0.28	0.25	0.44	0.46	0.45	0.48
Enrolled in Medicaid	0.20	0.17	0.42	0.42	0.42	0.41
Parent's education, < HS	0.31	0.29	0.47	0.48	0.48	0.49
Parent's education, HS	0.26	0.26	0.27	0.26	0.26	0.27
Parent's education, some college	0.27	0.28	0.19	0.18	0.18	0.17
Parent's education, college	0.13	0.14	0.06	0.06	0.07	0.06
Parent's education, missing	0.03	0.03	0.02	0.02	0.01	0.02
Total Observations	23,600	20,103	3,497	3,019	1,818	1,201

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.

Table 2.5: Summary Statistics of Applicants' Outcomes by Treatment Status

Variable	Cohort	All Applicants		Non-Semifinalists		Semifinalists		Finalists		Non-Scholars		Scholars	
		Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
Immediate enrollment	2009-2012	0.73	23,600	0.71	20,103	0.82	3,497	0.84	3,019	0.82	1,818	0.88	1,201
2nd year persistence rate	2009-2012	0.62	23,600	0.60	20,103	0.72	3,497	0.75	3,019	0.70	1,818	0.82	1,201
3rd year persistence rate	2009-2012	0.55	23,600	0.53	20,103	0.64	3,497	0.67	3,019	0.61	1,818	0.76	1,201
Bachelor's degree on time	2009-2012	0.20	23,600	0.19	20,103	0.25	3,497	0.27	3,019	0.22	1,818	0.34	1,201
Bachelor's degree in 6 years	2009-2010	0.49	10,252	0.47	8,556	0.58	1,696	0.60	1,454	0.53	853	0.70	601

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.

2.4.1 Validity of the Regression Discontinuity Design

We take advantage of the Dell Scholars Program selection processes to identify the causal impact of being selected as a scholar on college enrollment, persistence and completion outcomes. We exploit the fact that the program uses well-specified rank thresholds for the selection of scholars. Specifically, the program first ranks all applicants based on their final application scores and selects the top 900 scoring applicants as semifinalists. Then, among those semifinalists who complete the finalist application, the program scores and ranks finalists and selects the top 300 scoring finalists as Dell Scholars, such that, in each year, the score of the 300^{th} ranked finalist is the relevant threshold for examining the effect of being selected as a Dell Scholar.

These features allow us to use a regression discontinuity (RD) design to compare the outcomes of finalists with scores just above and below their year-relevant scholar-selection threshold. The students with scores just around the thresholds are comparable on many dimensions, however the finalists with scores just above the relevant thresholds were selected as the program's scholars. Thus, we can rely on the comparison of students at the scholar-selection margins to obtain unbiased estimates of the impacts of scholar selection.

For a regression discontinuity strategy to yield valid causal inference, several conditions must be met (Schochet et al., 2010). First, the assignment rule must be clear and followed with a high degree of fidelity. Second, the score utilized to determine scholar status, our running variable, should be an ordinal measure with sufficient density on either side of the cut off. Third, these scores should be utilized by the Dell Scholars program only for the purpose of identifying scholar status, and therefore differences that we see at the relevant margin cannot be attributable to other potential mechanisms. Finally, applicants must not able to manipulate their own value of the running variable. Regarding this final point, it is highly implausible that applicants would have this ability. The scoring algorithms are complex, are not publicly disclosed, and rely on multiple inputs. Further, manipulation of one's position relative to the cutoffs would require perfect information of the selection processes as well as of the inputs associated with other applicants.

An additional, related potential threat to validity is rater manipulation, such that raters are overly generous in scoring certain applications. While we cannot fully rule out the potential of rater manipulation, we argue that it is unlikely to have an undue influence on students' final rank order especially local to the threshold for scholar selection. First, each rater evaluates applications for only a small subset of all finalists. Second, the final scores are a combination of rater evaluation scores and scores attributed to measures like student GPA. In sum, raters are unlikely to know the marginal score that will be in the top 300 and are unlikely to be able to finely manipulate a student's overall score around that relevant margin. Finally, as noted in Section 2.2.3, raters are never assigned to review the applications of students who reside in the same zip code. Therefore, it is unlikely that raters have a personal connection with any of the students whose application materials they are reviewing.

2.4.1.1 Fidelity of the Assignment Rule We provide evidence that the remaining conditions are met in the context of the Dell Scholars program. In Figure 2.1, we illustrate the relationship between scholar status and the final score, by year. In each year, the relevant threshold is demarcated with a vertical dashed line. These figures provide evidence that the selection rules and processes are followed with a very high – almost perfect – degree of fidelity. Nevertheless, we do observe a few exceptions to the stated selection rules. We find a small number of instances where finalists, whose scores are above the scholar threshold, were not selected as scholars.

In Table 2.6, we report on the relevant threshold values for the identification of scholars and report counts of the number of cases in which the selection rules were not strictly

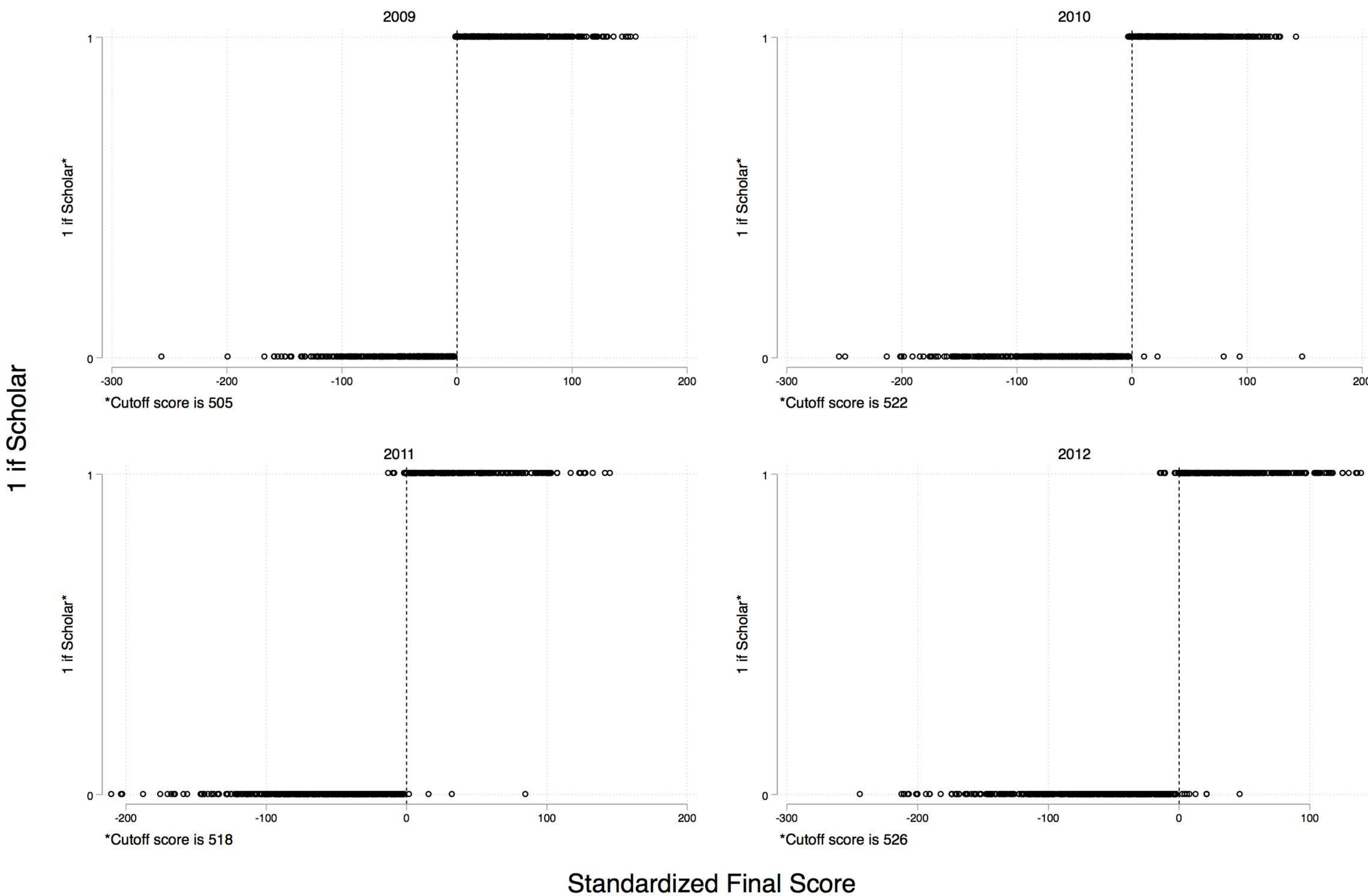

Figure 2.1: Relationship between Scholar Status and Final Score by Year. Source: The Dell Scholars Program database, Michael and Susan Dell Foundation.

followed, by year. These instances of non-compliance with the stated assignment rules are explained by the fact that the Dell Scholar team reserves the right to manually disqualify applicants after they initially have been selected as a scholar. There are four main reasons for disqualification. First, an applicant may be disqualified if the applicant received a serious disciplinary action in high school. The Dell Scholar Program has yet to disqualify a scholar for this reason. Second, an applicant may be disqualified if the applicant's essay did not meet the minimum criteria or if the applicant used the same responses for all essays. Third, an applicant may be disqualified if the applicant did not plan to attend a four-year college. While it is permissible for scholars to begin their postsecondary education at a community college, they must demonstrate a goal of completing a four-year degree. Lastly, an applicant may be disqualified if the applicant inflated their high school grades. Specifically, the Dell Scholar Program checks whether the self-reported grades matched with the official high-school transcript. Despite these small discrepancies, collectively we have strong evidence in support of an RD strategy for assessing programmatic impacts.

Table 2.6: Threshold Scores and Assignment of Scholars by Year

Cohort	Threshold Score	Non-Scholars with score below threshold	Non-Scholars with score above threshold	Scholars with score below threshold	Scholars with score above threshold	Scholars
2009	505	343	0	0	300	300
2010	522	505	5	6	295	301
2011	518	457	4	4	295	299
2012	526	496	8	8	293	301

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.

We formally estimate the relationship between scholar status and assignment rule. We use the following linear probability model for finalist i in cohort t:

$$SCHOLAR_{it} = \beta_0 + \beta_1 ASSIGN_{it} + \beta_2 SCORE_{it} + \beta_3 \left(ASSIGN \cdot SCORE\right)_{it} + \Gamma \mathbf{X}_{it} + \varepsilon_{it}, \tag{2.1}$$

where $SCHOLAR$ is an indicator for selection as a Dell Scholar, $ASSIGN$ is an indicator for a final score that is equal to or higher than the cohort-specific threshold, and $SCORE$ is the final score re-centered around the threshold. Note that under a perfect assignment rule, the variables $SCHOLAR$ and $ASSIGN$ are equivalent for all students. The vector $\mathbf{X}$ comprises control covariates including cohort dummies, interactions between cohort dummies and the running variable, interactions between cohort dummies, the running variable, and the scholar status, age, scaled GPA, ACT equivalent score, dummies for state of residence, a female dummy, ethnicity dummies, parents' adjusted gross income, dummies for parental education, free or reduced-lunch eligibility, receipt of food stamps, receipt of federal health insurance, receipt of Medicaid, an indicator for missingness of ACT, and an indicator for missingness of food stamp receipt. Note that our specification allows the slope of the relationship between the probability of being selected as a scholar and scores to vary above and below the cohort-specific thresholds.

The parameter of the assignment indicator, β_1, represents the difference in the probability of being selected as a scholar between students who are just above and just below the year-specific threshold. The parameter is equal to 1 if the assignment rule is followed perfectly. We expect the parameter to be less than but close to 1 since the assignment rule was followed with near perfect fidelity. Indeed, as shown in Column 1 of Table 2.7, we estimate that a finalist with a score just above the threshold has a 0.955 higher probability of being selected as a scholar. In the remaining columns, we present estimates associated with the intermediate, the narrow, and the optimal bandwidths. Across columns, the results are very similar in terms of both magnitude and statistical significance and are therefore not sensitive to bandwidth selection. The results show that, albeit imperfection in the assignment rule, there is a high-degree of fidelity. These results motivate our modeling strategy which we will discuss in Section 2.4.2.

2.4.1.2 Continuity of the Running Variables We now test the validity of the RD assumptions related to the continuity of the running variables across the relevant thresholds

Table 2.7: Relationship between Scholar Status and Assignment Rule

	1: Full sample	2: Intermediate	3: Narrow	4: Optimal
Assignment rule	0.955***	0.953***	0.947***	0.946***
	(0.010)	(0.009)	(0.011)	(0.012)
Score	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Score x assignment rule	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	3019	2992	2299	1944
Adjusted R^2	0.95	0.96	0.95	0.95
Mean below threshold	0.03	0.02	0.02	0.02

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.
Notes: Robust standard errors are in the parentheses and sample sizes are in brackets. The signs *, **, *** indicate significance at 10%, 5%, and 1% level, respectively. Explanatory variables not shown in the table are the running variable, an interaction between scholar status and the running variable, cohort dummies, interactions between cohort dummies and the running variable, interactions between cohort dummies, the running variable, and the scholar status, age, scaled GPA, ACT equivalent score, dummies for state of residence, a female dummy, ethnicity dummies, dummies for parental education, free or reduced-lunch eligibility, receipt of food stamps, receipt of federal health insurance, receipt of Medicaid, an indicator for missingness of ACT score, and an indicator for missingness of food stamp receipt.

and assess any evidence of manipulation of position around these thresholds. We utilize the McCrary (2008) test to examine the continuity assumption of the running variables by assessing the smoothness of the score densities across the relevant thresholds. The intuitive purpose of this exercise is to test whether applicants were able to manipulate their assignment. Graphically, we expect continuity in the density of the continuous assignment variable around the threshold if applicants were not able to manipulate their assignment. Given the complexity of the scoring algorithms utilized as well as the fact that semifinalist and scholar selection thresholds are determined relative to each year's pool of applicants (rather than being an absolute, pre-determined threshold), we expect for the assumptions to be met.

In Figure 2.2, we illustrate the graphical presentation of the McCrary test for the finalists' scores. In Panel A of Figure 2.2, we present results by year and in Panel B of Figure 2.2, we present data pooled across the 2009 through 2014 cohorts. In each panel, we can observe the continuity of the assignment variable. In Table 2.8, we provide summary statistics from the McCrary tests. In all but one cohort and for results pooled across cohorts, the test fails to reject the null hypothesis at the 5 percent significance level. The only instance where the data failed to reject the null hypothesis is in the 2011 scholar selection. This result may be

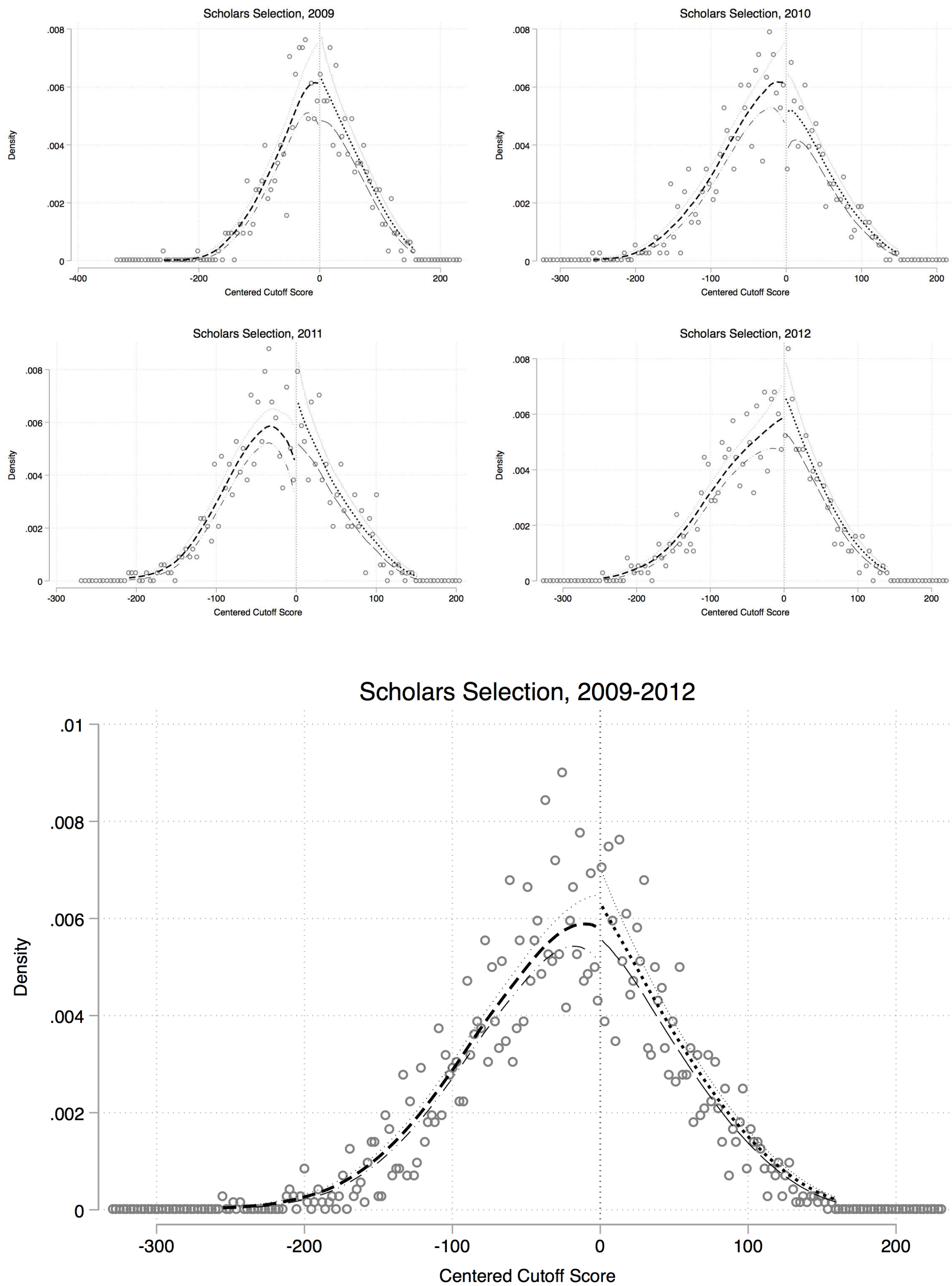

Figure 2.2: Graphical Representation of the McCrary Density Tests by Year for the Scholar Selection. Source: The Dell Scholars Program database, Michael and Susan Dell Foundation.

driven by the missing data of 138 finalists and 1 scholar in the 2011 finalist dataset.[9]

Table 2.8: Results of the McCrary Density Tests by Cohort

Cohort	**Finalists**		
	Mean	**S.E.**	t**−statistics**
2009	0.049	0.180	0.275
2010	-0.177	0.185	-0.954
2011	0.460	0.205	2.250
2012	0.128	0.156	0.822
2013	-0.320	0.173	-1.850
2014	-0.210	0.195	-1.073
2009-2014 (pooled)	-0.019	0.076	-0.243

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.

As an additional check, we examine potential jumps in finalists' characteristics at the scholar selection thresholds. Finalists' characteristics that we analyze include gender, age, ACT equivalent score, high-school GPA, ethnicity, enrollment in lunch program, and first-generation college student status. We conduct a graphical analysis before we fit regression discontinuity models. First, we implement a bandwidth selector proposed in Calonico et al. (2014a) to obtain an optimal bandwidth for each variable. We use the *rdbwselect* package in STATA that implements this bandwidth selector (Calonico et al., 2014b). We specify a local linear regression with a uniform kernel for the bandwidth selector because we fit regression discontinuity models using OLS.[10] Second, we use the *rdplot* package in STATA which implements data-driven RD plots (Calonico et al., 2015).

We present the RD plots in Figure 2.3. The graphical analysis suggests no jumps around the threshold for the gender, age, and ACT equivalent score variables. However, we can observe slight jumps for the remaining variables. To test whether these jumps are significantly different from zero, we fit regression models. Specifically, we use a linear regression model for finalist i in cohort t:

$$C_{it} = \alpha_0 + \alpha_1 SCHOLAR_{it} + \alpha_2 SCORE_{it} + \alpha_3 \left(ASSIGN \cdot SCORE\right)_{it} + \Gamma \mathbf{X}_{it} + \epsilon_{it} \quad (2.2)$$

[9] We assess the sensitivity of our results to this cohort and find that results are, overall, not sensitive to the inclusion or exclusion of the class of 2011 students for who score data are complete.

[10] See Lee and Lemieux (2010) and Imbens and Lemieux (2008) for further discussions.

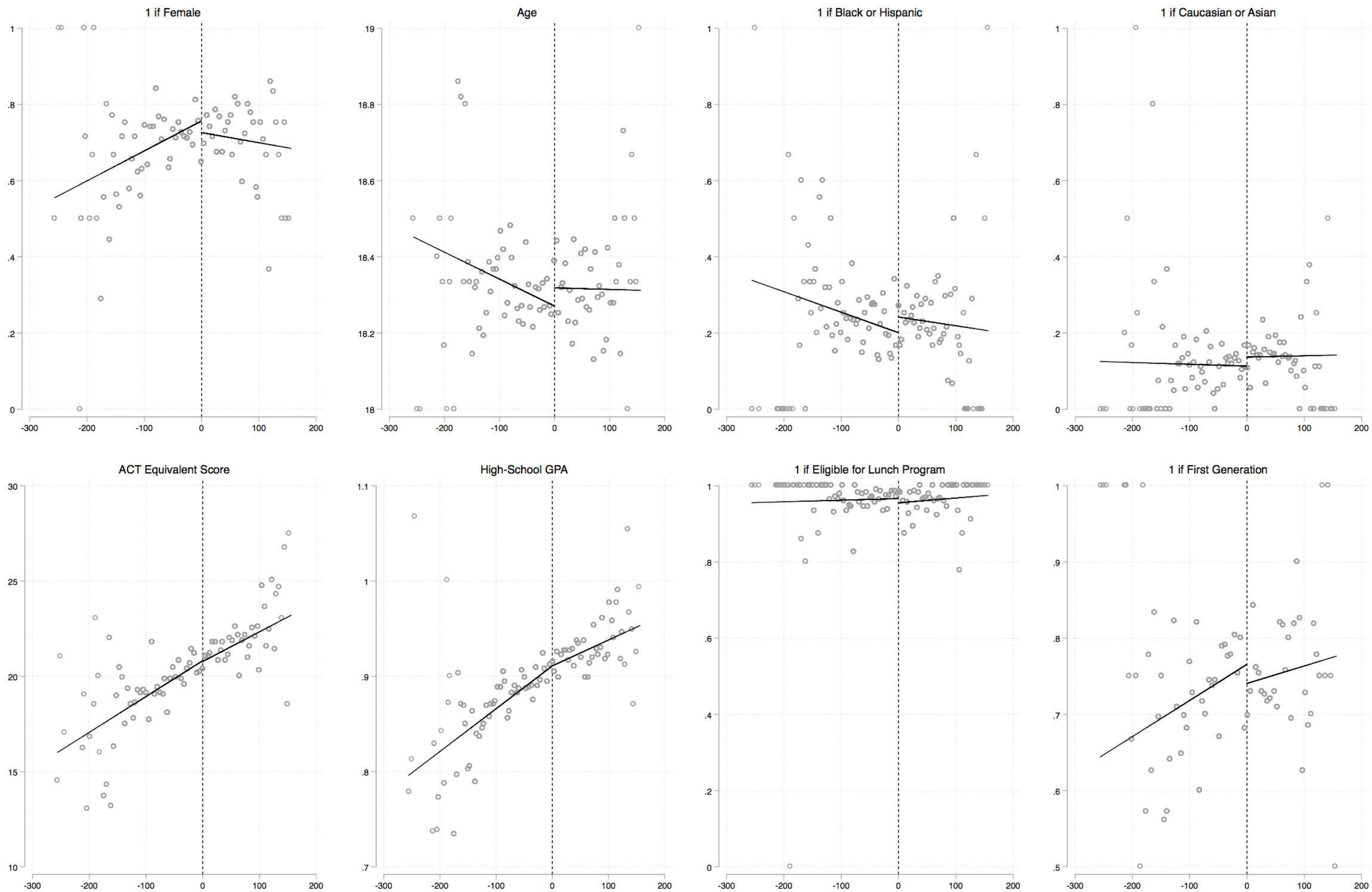

Figure 2.3: RD Plots of Finalists' Characteristics. Source: The Dell Scholars Program database, Michael and Susan Dell Foundation.

where C represents a finalists' characteristic. The variable $SCHOLAR$ takes a value of 1 if the finalist received a Dell Scholarship. The variable $ASSIGN$ takes a value of 1 if the finalist's final score is above the cohort-specific cutoff score. The running variable, $SCORE$, is the centered finalist scores. Finally, the vector $\mathbf{X}$ includes the same set of covariates as the ones in Specification 2.1. We exclude a variable from the vector $\mathbf{X}$ if the variable is used as the left-hand side variable (C).

We present the regression results in Table 2.9. In no case did these analyses reveal that the jumps around the selection cutoff are significantly different from zero. Collectively, the results in this section do not point to evidence of manipulation of the running variables. We conclude that the Dell Scholar selection rules generate a robust quasi-random assignment of scholars local to the relevant selection thresholds.

Table 2.9: Relationship between Covariates and Scholar Status Using the Full Sample

Finalists' characteristics	**1 if Female**	**Age**	**1 if Black or Hispanics**	**1 if White or Asian**
1 if Scholar	-0.026	0.006	-0.001	-0.001
	(0.023)	(0.026)	(0.004)	(0.003)
Total observations	4,605	4,605	4,605	4,605
Finalists' characteristics	**ACT equivalent**	**High-school GPA**	**1 if lunch program**	**1 if first-generation**
1 if Scholar	0.200	-0.016	0.003	0.012
	(0.168)	(0.018)	(0.008)	(0.020)
Total observations	4,605	4,605	4,605	4,605

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.
Notes: Robust standard errors are in the parentheses and sample sizes are in brackets. The signs *, **, *** indicate significance at 10%, 5%, and 1% level, respectively. Explanatory variables not shown in the table are the running variable, an interaction between scholar status and the running variable, cohort dummies, interactions between cohort dummies and the running variable, interactions between cohort dummies, the running variable, and the scholar status, age, scaled GPA, ACT equivalent score, dummies for state of residence, a female dummy, ethnicity dummies, dummies for parental education, free or reduced-lunch eligibility, receipt of food stamps, receipt of federal health insurance, receipt of Medicaid, an indicator for missingness of ACT score, and an indicator for missingness of food stamp receipt

2.4.2 Model Specification

The modest infidelity to the selection processes suggests that we need to utilize a two-stage instrumental variables (IV) or fuzzy RD approach (Jacob and Lefgren, 2004; Imbens and Lemieux, 2008). Based on this approach, the assignment rule is an instrument for the scholar selection. However, we have shown that the assignment rule is almost a perfect predictor of scholar selection. Therefore, we opt to use a reduced-form specification to estimate the impact of being selected as a Dell Scholar on college-going outcomes.[11] Specifically, we use a linear probability model for finalist i in cohort t:

$$Y_{it} = \pi_0 + \pi_1 ASSIGN_{it} + \pi_2 SCORE_{it} + \pi_3 \left(SCORE \cdot ASSIGN\right)_{it} + \theta \mathbf{X}_{it} + \tau_{it}, \qquad (2.3)$$

where Y represents a particular outcome such as college enrollment, persistence, or bachelor degree completion. The parameter π_1 indicates the causal impact of being selected as a Dell Scholar. We use the same set of covariates as the ones in Specification 2.1. We estimate the parameters of the models using different bandwidths: the full sample, an intermediate bandwidth of ± 100 points, a narrow bandwidth of ± 40 points, and an optimal bandwidth around the threshold.[12]

In selecting the optimal bandwidth, we utilize a first-order polynomial, a uniform kernel, and the bandwidth selector of Calonico et al. (2014a). The optimal bandwidth varies across outcomes because of this process. For this reason, we choose the two additional bandwidths of ± 100 and ± 40 points which are constant across outcomes. We estimate the impacts of being selected as a Dell Scholar on different outcomes among students in the 2009-2012 cohorts to avoid composition effects. Note that the only students we are able to observe for a full six year of college enrollment are students in the 2009-2010 cohorts. Therefore, we also estimate the impacts among students in the 2009-2010 cohorts.

[11] We find that fuzzy RD estimates are quite identical to reduced-form estimates.

[12] Lee and Lemieux (2010) suggest estimation of linear regression model using different bandwidths as opposed to estimation of linear regression model with different kernels.

We present impacts of being Dell Scholars on the college-going outcomes in Table 2.10.[13] We estimate that students just below the threshold have a 0.85 probability of enrolling in college. At the margin of selection, being selected as a Dell Scholar improves timely enrollment in 4-year institutions among scholars in the 2009-2010 cohorts. We estimate that being selected as a Dell Scholar improves enrollment by about 9.8 percentage points. However, we find no significant impact of being selected as a Dell Scholar on timely enrollment when we include the 2011 and 2012 cohorts. The lack of impact with respect to college enrollment is not necessarily surprising. As discussed above, all students who achieve finalist status in the Dell Scholars application process are likely to be highly college intending, and students were notified of their scholar status well after deciding where to apply and, for many, where to attend. Even for these highly college-intending students, however, it is notable that a sizable share is not successfully matriculating to college, potentially facing other barriers to timely postsecondary enrollment (e.g. Castleman and Page, 2014b,a).

The generous financial support that Dell Scholars receive may substantially alleviate the financial constraints that students and families experience in covering costs associated with college attendance. These include academic costs, such as tuition, fees, and books, as well as non-academic costs, such as child care. If the scholarship improves students' ability to finance college, semester over semester, then we may expect to see substantially higher persistence among Dell Scholars. Second, through gathering data on scholars' postsecondary experiences and providing them with feedbacks and supports, as needed, the program may provide students with the guideposts, encouragement, and direction that they need to be more successful throughout their college careers. Improvements in outcomes such as persistence as well as other success metrics such as college GPA, number of credits attempted, and

[13] We present the fuzzy RD estimates in Table B.1. The fuzzy RD estimates are quite identical to reduced-form estimates.

number of credits earned each semester would align with these mechanisms.

While we lack data for both scholars and applicants on academic performance in college, we do utilize data from the National Student Clearinghouse to examine year-by-year persistence outcomes through the first three years of college. We present these results in Table 2.10. Like immediate enrollment, we find no impact on second-year persistence at the margin of being selected as a scholar. However, we find a strong and positive impact of being a Dell Scholar on third-year persistence. We estimate that students just below the threshold have a 0.67 probability of persisting to the third year while students who just meet the threshold for scholar selection are 5-11 p.p. more likely to persist to the third year. The impact on third-year persistence is even higher among scholars in the 2009-2010 cohort as those who just meet the threshold for scholar selection are 7-18 p.p. more likely to persist to the third year. We interpret these results as a strong evidence that the program has a positive impact on postsecondary persistence, especially in the later years of college.

We also find positive and statistically significant impacts of being a Dell Scholar on degree completion. Among the 2009-2012 cohorts, the estimated probability of obtaining a bachelor's degree within 4 years is about 0.29. This estimate suggests that about 43 percent of finalists who persisted to the third year earned a bachelor's degree within 4 years. Our estimate suggests that students who just meet the threshold for scholar selection are about 6 p.p. more likely to complete a bachelor's degree in 4 years. The estimated impacts are larger among scholars in the 2009-2010 cohorts. Specifically, students who just meet the threshold for scholar selection are 8-10 p.p. more likely to earn a bachelor degree on time, and 16 p.p. more likely to earn a bachelor degree in 6 years.[14] The impacts that we estimate are visually apparent in Figure 2.4 which illustrates the impacts of being selected as a Dell Scholar on immediate enrollment, third-year persistence, on-time BA attainment, and BA attainment in 6 years, respectively.[15]

[14] We observe in the data that almost all students who earned a bachelor degree do so within 6 years.

[15] As suggested by Lee and Lemieux (2010), we estimate the model using different bandwidths around the threshold. We present the results of estimations using different optimal bandwidths in Table B.2 for the 2009-2012 cohorts and in Table B.3 for the 2009-2010 cohorts. We find that the magnitudes of the effects

Table 2.10: Impacts of Scholar Selection on College-Going Outcomes

Outcomes	2009-2012 Cohorts						2009-2010 Cohorts					
	μ	Full Sample	Intermediate ($\pm$100)	Narrow ($\pm$40)	Optimal Bandwidth	Range of Bandwidths	μ	Full Sample	Intermediate ($\pm$100)	Narrow ($\pm$40)	Optimal Bandwidth	Range of Bandwidths
1: Intermediate enrollment	0.853	0.015	0.022	0.023	0.014	44	0.850	0.032	0.037	0.064	0.098*	36
		(0.022)	(0.025)	(0.037)	(0.036)			(0.031)	(0.036)	(0.054)	(0.057)	
		[3,019]	[2,585]	[1,372]	[1,482]			[1,454]	[1,245]	[668]	[602]	
2: 2^{nd} year persistence rate	0.752	0.023	0.036	0.032	0.025	44	0.751	0.034	0.054	0.072	0.103	38
		(0.026)	(0.030)	(0.044)	(0.043)			(0.037)	(0.043)	(0.064)	(0.067)	
		[3,019]	[2,585]	[1,372]	[1,482]			[1,454]	[1,245]	[668]	[631]	
3: 3^{rd} year persistence rate	0.674	0.046*	0.060*	0.081*	0.110**	36	0.658	0.067*	0.090*	0.121*	0.184**	36
		(0.028)	(0.032)	(0.048)	(0.050)			(0.040)	(0.046)	(0.069)	(0.074)	
		[3,019]	[2,585]	[1,372]	[1,240]			[1,454]	[1,245]	[668]	[602]	
4: BA attainment, in 4 years	0.287	0.062**	0.056*	0.065	0.059	34	0.266	0.081**	0.096**	0.101	0.090	31
		(0.028)	(0.032)	(0.051)	(0.055)			(0.041)	(0.047)	(0.075)	(0.087)	
		[3,019]	[2,585]	[1,372]	[1,176]			[1,454]	[1,245]	[668]	[530]	
5: BA attainment, in 6 years							0.633	0.042	0.056	0.093	0.158*	30
								(0.042)	(0.048)	(0.075)	(0.090)	
								[1,454]	[1,245]	[668]	[512]	

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.
Notes: Robust standard errors are in the parentheses and sample sizes are in brackets. The signs *, **, *** indicate significance at 10%, 5%, and 1% level, respectively. Explanatory variables not shown in the table are the running variable, an interaction between scholar status and the running variable, cohort dummies, interactions between cohort dummies and the running variable, interactions between cohort dummies, the running variable, and the scholar status, age, scaled GPA, ACT equivalent score, dummies for state of residence, a female dummy, ethnicity dummies, dummies for parental education, parents' income, free or reduced-lunch eligibility, receipt of food stamps, receipt of federal health insurance, receipt of Medicaid, an indicator for missingness of ACT score, and an indicator for missingness of food stamp receipt. To obtain the optimal bandwidth, we use a first-order polynomial, a uniform kernel, and bandwidth selector of Calonico et al. (2014a).

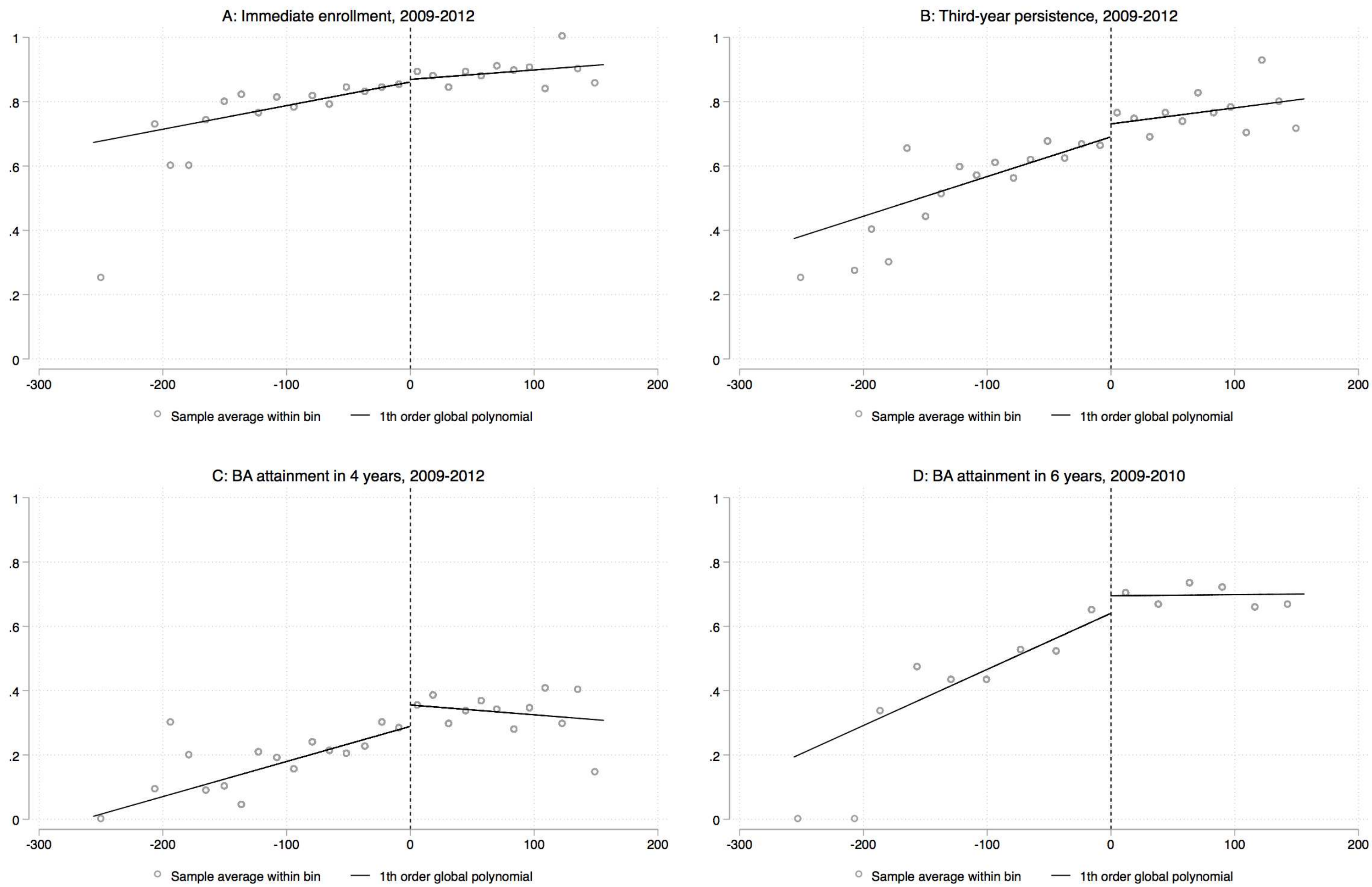

Figure 2.4: Regression Discontinuity Plots of Scholar Selection and College-Going Outcomes. Source: The Dell Scholars Program database, Michael and Susan Dell Foundation.

As is typical in the context of RD analysis, we face a trade-off between statistical power and estimating treatment effects local to the relevant thresholds (e.g. Ludwig and Miller, 2007). For instance, in Table 2.10, the magnitude of on-time degree completion effects is quite consistent across the four different bandwidths while statistical significance differs in some instances due to a loss of precision when restricting the sample. Nevertheless, we judge these results collectively to provide compelling evidence regarding the substantial impact of the Dell Scholar opportunity on college completion. Specifically, these impacts represent about an 11 percent increase in third-year persistence and a nearly 22 percent increase in on-time bachelor's degree attainment.

Given these sizable impacts on postsecondary persistence and degree attainment, a key question is what are the mechanisms by which the program is improving persistence and college completion. One plausible explanation is that being selected as a Dell Scholar may impact the type or "quality" of the institution in which students enroll. This may occur for two reasons. Upon being selected as a Dell Scholar, students participate in in-take interviews conducted prior to postsecondary matriculation. During this interview, the Dell Scholar team provides feedbacks on college plans and, in some cases, counsels students about certain postsecondary choices, such as planning to enroll in public institutions. The justification here is that a public institution is likely to be a more financially viable option. Second, the Dell Scholar award includes a sizable amount of grant-based financial aid that may enable students to view a different set of postsecondary options as within reach financially. If students are enrolling in "higher quality" institutions as a result of the Dell support, this may translate to better college completion outcomes (Goodman et al., 2015; Hurwitz et al., 2016; Howell and Pender, 2016)

We examine a set of indicators related to institutional quality and type. Related to quality, we specifically examined whether students at the margin of scholar selection initially enrolled in institutions that differed in terms of institutional graduation rates and in terms

are not sensitive to changes in the optimal bandwidth.

of instructional expenditure, given evidence from Hoxby et al. (2013) that instructional expenditure correlates with other metrics of institutional quality. We also examine whether students at the margin of the threshold for scholar selection initially enrolled in more competitive institutions. Related to college type, we investigate the impact of being selected as a Dell Scholar on whether they enrolled in public or private institutions. We present the RD estimates for these outcomes in Table 2.11. It is important to note that about 15 percent of the finalists have missing quality indicators because their institutions are not observed in the integrated postsecondary education data system (IPEDS). Therefore, we investigate whether there is a systematic relationship between being selected as a Dell Scholar and missingness. As shown in Row 1 of Table 2.11, we find no evidence that missingness are correlated with being selected as a Dell Scholar.

We find systematic evidence that students' specific institutional choices change along the quality dimensions as a result of being selected as Dell Scholars. We estimate that students who just meet the threshold of scholar selection initially enrolled in school with significantly higher instructional expenditure per full-time equivalent (PFTE) students. We also estimate that students who just meet the threshold of scholar selection are 6-7 p.p. more likely to enroll in either very, highly, or most competitive college. Overall, these results support our view that being selected as a Dell Scholar induces students to choose better quality schools, which in turn translates to higher persistence and completion rates.

2.6. DISCUSSION

We find compelling evidence of strong and significant impacts on later persistence and significant impacts on college completion. At the margin of being selected into the Dell Scholar Program, scholars are 5-18 percentage points more likely to persist to the third year, 6-10 percentage points more likely to earn a bachelor's degree on-time, and nearly 16 percentage points more likely to do so within six years than they would have been absent the Dell

Table 2.11: Impact of Scholar Selection on College Quality and College Type, 2009-2012 Cohorts

Outcomes	μ	Full Sample	Intermediate ($\pm$100)	Narrow ($\pm$40)	Optimal Bandwidth	Range of Bandwidths
1: 1 if missing quality indicators	0.150	0.020	-0.006	-0.006	-0.006	30
		(0.023)	(0.026)	(0.041)	(0.047)	
		[3,019]	[2,585]	[1,372]	[1,049]	
2: Institutional degree completion PFTE students	24.42	1.002*	0.645	0.250	0.429	26
		(0.519)	(0.586)	(0.815)	(0.942)	
		[3,019]	[2,585]	[1,159]	[1,482]	
3: Instructional expenditure PFTE students	$4,590	436.8*	609.8**	867.6**	550.8	36
		(265.3)	(293.4)	(433.3)	(557.9)	
		[2,556]	[2,202]	[1,159]	[770]	
4: Barron's category of very, highly, or most competitive college	0.487	0.062**	0.073**	0.003	-0.047	34
		(0.030)	(0.035)	(0.054)	(0.070)	
		[2,556]	[2,202]	[1,159]	[770]	
5: Enrollment in 4-year public institution	0.603	0.025	0.008	0.076	0.076	31
		(0.032)	(0.038)	(0.059)	(0.068)	
		[2,556]	[2,202]	[1,159]	[921]	

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation, and IPEDS Analytics: the Delta Cost Project Database.
Notes: Robust standard errors are in the parentheses, sample size in brackets. The signs *, **, *** indicate significance at 10%, 5%, and 1% level, respectively. Explanatory variables not shown in the table are the running variable, an interaction between scholar status and the running variable, cohort dummies, interactions between cohort dummies and the running variable, interactions between cohort dummies, the running variable, and the scholar status, age, scaled GPA, ACT equivalent score, dummies for state of residence, a female dummy, ethnicity dummies, dummies for parental education, free or reduced-lunch eligibility, receipt of food stamps, receipt of federal health insurance, receipt of Medicaid, an indicator for missingness of ACT score, and an indicator for missingness of food stamp receipt.

Scholar opportunity.

The magnitude of the degree effects are similar in structure to those observed by Scott-Clayton (2011) in her examination of the West Virginia (WV) PROMISE program (this paper also utilizes a regression discontinuity design with an academically-similar population of students). Scott-Clayton concludes that the WV PROMISE program provided academic achievement "guideposts" to students that helped to improve the quality of their postsecondary enrollment. Students supported through the WV PROMISE earned more credits over four years and achieved higher GPAs, for example, than their non-PROMISE counterparts, due to the motivation of the structural requirements of the financial support. Although we are not able to observe these "process" measures in the data to which we currently have access, we hypothesize that the Dell opportunity may operate through improving the academic success of the Dell Scholars by combining generous financial support with both guideposts for success and close monitoring, feedback and support to keep scholars on track for degree attainment.

The impact that we observe on on-time bachelor's degree attainment (6-10 percentage points) is also of similar magnitude to that observed by Scott-Clayton at the margin of PROMISE eligibility. In that only a third of the WV students qualified for a Pell Grant compared to all of the Dell Scholar recipients, the magnitude of the Dell Scholars program impact on on-time degree attainment is particularly remarkable. Further, different from the WV PROMISE context, we observe that over a longer time horizon, the impact on degree completion grows. Specifically, we estimate a 16 percentage point impact on attainment of a bachelor's degree within six years.

Given the positive impacts of the Dell Scholars Program on degree attainment, an important question is whether the benefits associated with these increases in college completion justify the costs of the program. Therefore, we provide a back-of-the envelope calculation regarding the relative costs and benefits of the program in the spirit of Deming (2009), Pallais (2015), and Hurwitz et al. (2016). We consider about 1,000 students who were close to the threshold and to whom the estimated impact would apply. Our estimate using the optimal bandwidth suggests that 158 more students (or 15.8 percent) would earn a bachelor's degree. The total costs of the Dell Scholars Program for these students is $30 million or $30,000 per student. This implies that the cost to induce one additional student to earn a bachelor's degree is about $189,837.

We now consider the benefits of the higher attainment of a bachelor's degree. It is estimated that in 2011 median full-time workers with a bachelor's degree earned, on average, $16,100 more in annual earnings and tax payments than full-time workers with only some college (Baum et al., 2013). While this is an observed difference, Card (1999) reports that causal estimates of the schooling effect earnings are often 20 to 40 percent larger. Assuming a constant earnings differential between the two groups, the Dell Scholar Program would reap a net benefit after twelve years of post-college earning. Even if the earnings differential between Dell Scholars recipients and finalists just below the threshold were smaller, given that the latter group persisted through several years of college at more similar rates, the program is still expected to have a positive rate of return, albeit over a long time horizon.

Of course, this simple calculation leaves aside many factors. For example, we might consider this estimate conservative, in that we do not attempt to monetize the many other types of benefits, both public and private, that accrue as a result of higher education (Baum et al., 2013). Similarly, we do not adjust for an increase in earnings differentials over time. While recognizing the many assumptions that we have made, these calculations nevertheless suggest a positive rate of return for the Dell investment in their Dell Scholars program.

Some college access and persistence efforts focus on financial barriers to college success by providing students with scholarship funds. Other efforts focus on additional outreach and counseling to assist students in navigating the academic and behavioral challenges that emerge in college. While evidence suggests that both types of efforts hold promise for improving the college-going outcomes of low-income and first-generation college-going students, it may be that offering students a suite of supports across these domains may be more successful than the sum of its parts. The ASAP program in New York City suggests this to be true in the community college context. Our examination of the Dell Scholars program provides further supporting evidence, primarily in the context of four-year colleges and universities.

As discussed above, there are several eligibility criteria that students must meet to initially apply and be selected as a Dell Scholar. Upon selection, there are several steps that students must take in order to remain eligible. This includes regular reporting back to the Dell Scholars team on academic progress as well as challenges that they are facing, be they related to academics, physical health, mental health, college finances, or general life management. By incorporating this reporting mechanism into their ongoing work with scholars, the Dell Scholars team is able to track their students closely and triage additional team support to students when needed.

Our results indicate that this support, coupled with generous and flexible financial aid, leads to improved rates later persistence, on-time and within six years of college completion. Taken together, the results point to the Dell Scholar program supporting scholars to be more efficient and effective in their postsecondary educational experiences. Although we are not able to shed light on the specific mechanisms through which the program operating, in subsequent work, we will turn to a rich investigation of the process through which the Dell Scholars program helps students to persist and succeed in college through to degree completion. This work will help to inform the college access and success community in efforts to go beyond initial college enrollment to focus on ultimate degree attainment and to understand the many facets of the college experience with which students may benefit from increased structure, guidance and support.

3. THE PRICE OF RELIGION: EXPERIMENTS IN WILLINGNESS TO BEAR RISK FOR OTHERS IN ISLAMIC COMMUNITIES

Co-authored with Sera Linardi, Rebecca B. Morton, Kai Ou, and Xiangdong Qin

3.1. INTRODUCTION

Leaders of social movements often appeal to religion.[1] Participation in these movements costs not only time but it also costs money and effort, such as pledging a donation, signing a petition, or participating in a demonstration. Many times, individuals who join these movements impose risks on themselves in order to achieve a prosocial outcome. For instance, individuals join a disaster relief team to help disaster victims. The degree to which individuals are willing to take these risks in this context are different from one individual to another. Some individuals are willing to take risks only when the costs imposed to themselves are low. On the other hand, there are individuals who are willing to take risks even when the costs imposed to themselves are high. In this study we ask whether using a religious precept is an effective way to motivate individuals to take prosocial risks particularly when the costs imposed to themselves are high?

This paper explores the effect of a religious message on prosocial risk-taking. We investigate this effect in the context of a question that is important in Islam but understudied

[1]This research is funded through the New York University Abu Dhabi Research Enhancement Fund. We are grateful for comments and suggestions from participants of the Pitt Experimental Economics Seminar, the 2015 ASREC Conference, and the 2015 New York University CESS Experimental Political Science Conference.

in the Western world: The Islamic prohibition against lending with interest. Specifically, Islam prohibits lending with interest-based arrangements (IB) and promotes profit- and loss-sharing arrangements (PLS).[2] The main difference between PLS and IB arrangements is the risk-sharing feature.[3] A lender makes a relatively risk-free investment in an IB arrangement since a borrower must pay back the loan and the interest whether he or she makes profits or not. However, a lender shares the risk of financing in a PLS arrangement. A lender obtains a return from investment only when the borrower makes profits, and makes a loss when the borrower makes a loss.

In our experiment, lenders, who are randomly matched to borrowers, choose whether to use a safer option in IB or a risky one in PLS. The lenders make decisions for two ranges of interest rates, 0-30 percent and 30-60 percent. We randomize whether before making the decision the lender receives a religious message on the Islamic prohibition against interest. The religious message we use is a verse in the Qur'an. We use a direct quote because it is quite common for leaders or organizations to cite Quranic verses during campaigns or movements to promote prosocial concepts such as justice or equality.

We focus our investigation on the interest rate range of 30-60 percent which provides a direct test of the effect of a religious message on the willingness to take risk to help others. First, taking up PLS is risky but it is the more prosocial choice as it eases the borrower's burden. Specifically, the borrower's expected payoffs under IB are strictly lower than the borrower's expected payoffs under PLS because under IB the borrower must pay back the loan and the interest even in the event of a loss. Second, this range of interest rates is policy relevant particularly in the microfinance context. A Consultative Group to Assist the Poor (CGAP) report by Rosenberg et al. (2013) shows that the global median interest yield is about 27 percent and that the interest yield can be as high as 70 percent.

We highlight two features of our experimental design. First, we develop a new method,

[2] There are two forms of PLS arrangements, *Mudarabah* and *Musharakah*. In this study, we focus on the *Mudarabah* arrangement.

[3] See Dar and Presley (2000) and Farooq (2007) for a discussion of PLS arrangement.

a bi-directional BDM, which allows us to investigate two prosocial behaviors. Lenders can appear prosocial by choosing PLS when the cost to them in terms of lost payoff is low. We label such lenders who are prosocial when costs to them are low as *protecting myself.* On the other hand, lenders can be prosocial when the costs are high. We label lenders who are prosocial when the costs are high as *protecting others.* Second, we provide the lenders with a "mouse-lab" calculator, which allows us to trace a lender's thought process. This feature allows us to see whether the religious message changes the extent to which the lender observes her own payoff or the borrower's payoff, and whether the lender observes the payoff of a successful or the payoff of an unsuccessful project.

Our study is large scale across two countries with quite distinct Muslim communities: Yogyakarta (Indonesia) and Ningxia Hui Autonomous Region (China). We also conduct experiments in Shanghai (China) to test whether the religious message is simply a message about a societal norm or a message about religious precept. Overall, we examine 1,328 decisions made by 332 subjects who express a wide range of degrees of religiosity. The cross-country setting allows us to investigate the differences in the effects of the religious message. More importantly, the setting allows us to speak about the external validity of our results and findings.

We find a higher proportion of *protecting-other* lenders in the message treatment. Specifically, there is a higher proportion of lenders who choose the risky PLS for high interest rates in the message treatment. Since lenders forgo relatively higher payoffs by choosing PLS for high interest rates, the lenders' expected gains are lower in the message treatment. We find that the effect of the religious message on expected gains are stronger among the more religious lenders. Our analysis of lenders' usages of the calculator show that the lenders in the message treatment use the calculator as frequently as the lenders in the baseline treatment. This result indicates that the lenders make conscious and informed choices to take risks for the borrowers after observing the religious message.

The religious message may simply be a message about morality. We conduct additional sessions in Shanghai to test whether the religious message convey a religious precept or a moral precept. From these sessions we find no difference between the take up rate of PLS among the lenders in the message treatment and the take up rate of PLS among the lenders in the baseline treatment. We conclude that the religious message is a message about a religious precept.

Our experiments on the 0-30 percent interest rate allow us to investigate whether the positive effects of the religious message is due to conformity or general altruism. In this range, choice of IB benefits the borrowers if the interest rate is low. Lenders face a conflict between choosing an option that conforms to the religious message and choosing an option that benefits the borrowers. We find no significant effect of the religious message on choices, which suggests that the effects is due to altruistic motives.

Our study complements findings that a religious priming motivates individuals to be more prosocial (Shariff et al., 2016). Our study also complements previous studies about the positive effects of religious priming on giving. Shariff and Norenzayan (2007), Sachs (2009), and Lambarraa and Riener (2015) find that individuals who are primed give more. Condra et al. (2017) finds that a direct use of a religious precept is significant among the religious. Specifically, they find that a quote from the Qur'an brings back the intrinsic motivation to donate from people whose intrinsic motivation had been crowded out. However, the design of these studies do not capture an important aspect of a prosocial behavior where people take risks to protect others.

The closest studies to ours in the prosocial literature are studies that investigate giving in risky environments. Two studies by Krawczyk and Le Lec (2010) and Brock et al. (2013) find lower amounts of transfer in a risky dictator game. A more recent paper by Exley (2016) finds that risk associated with the impact of giving induces excuses not to give, particularly when there is a trade-off between own payoff and recipient's payoff.

Our study is also related to a previous survey and an experiment that investigate individuals' choices between IB and PLS arrangements. El Massah and Al-Sayed (2013) study the relationship between risk aversion, credit experience, religion, and political views of 110 Egyptian subjects and hypothetical financing choices. After surveying subjects on the variables of interest, they present all subjects first with a hypothetical interest-bearing arrangement and ask whether they would accept that arrangement and then, afterwards, all subjects are presented with a hypothetical PLS arrangement and asked the same question. They find some evidence that subjects inexperienced with investments and higher degrees of religiosity are more willing to hypothetically accept the PLS arrangement. El-Komi and Croson (2013) conduct an experiment to test the compliance rate for IB and PLS arrangements under information asymmetry and costly state verification. In their study, individuals were assigned to either IB or PLS arrangement exogenously. They did not investigate choices between financing arrangements.

Lastly, our study is related to a study that investigate the effects of religious priming on risk-taking behavior Benjamin et al. (2016). They find that the religious priming induces individuals to become less risk averse on a gambling task. Our study is also related to surveys and experimental studies that investigate the relationship between risk attitudes and religiosity. Miller and Hoffmann (1995), Noussair et al. (2013), and Dohmen et al. (2011) find that more religious individuals are associated with a higher degree of risk aversion.

In the next section we describe the design of the experiment. In Section 3.3 we discuss the behavioral predictions. We discuss the results and the robustness checks in Section 3.4 and Section 3.5, respectively. Lastly, we discuss our conclusion in Section 3.6.

3.2. EXPERIMENTAL DESIGN

In our experiment subjects are randomly assigned a role as either lenders or borrowers and played an investment game. Lenders received an endowment of 30 points while the borrowers received one of 20 points. Lenders were required to loan 10 points of their endowment to the matched borrowers, and the borrowers paid the 10 points to the experimenter to complete a project. The borrowers received 20 points if the borrowers completed the project successfully, but the borrowers only received 6 points if the borrowers completed the project unsuccessfully.

The project consisted of answering a binary choice factual question such as "What is the biggest Island in Indonesia: (A) Papua or (B) Kalimantan." Such project does not require subjects to exert effort and thus avoids issues of moral hazard. The questions used were a combination of general knowledge questions, country-specific questions, and math questions. In each country we conducted several pretest sessions involving different subjects from the pool and selected questions such that the percentage of correct answers is approximately 67 percent.

The lender's choice of lending arrangement determined the lender's and the borrower's net payoffs. If the lender chose the IB arrangement, the borrower had to repay the loaned 10 points plus the interest, regardless of the outcome of the project. The borrower had to use the 20-point endowment to repay the loan and the interest if the project is unsuccessful. Given the expected success rate of 67 percent, the expected payoff of the lender was $10r$ while the expected payoff of the borrower was $5\frac{1}{3} - 10r$.

If the lenders chose the PLS arrangement, the lenders' and the borrowers' payoffs depended upon the success of the project. If the project is successful, the borrower shared the net profit (the 20-point gross profit minus the loaned 10 points) equally with the lender.[4] If the project is unsuccessful, the lenders only received 6 points while the borrowers received nothing. The lender's expected payoff was 2 points while the borrower's expected payoff was $3\frac{1}{3}$.

Figure 3.1 depicts the expected payoffs to a risk-neutral lender and borrower for interest rates between 30-60 percent. We can observe from the figure that the lender's expected payoff under IB is always higher than the lender's expected payoff under PLS. The difference in the expected payoffs between the two arrangements increases linearly with interest rate.

[4]The profit share of 50 percent is common knowledge to both lenders and borrowers.

Conversely, the borrower's expected payoff under IB is always lower than the borrower's expected payoff under PLS, and the difference in the expected payoffs between the two arrangements are decreasing with interest rate.

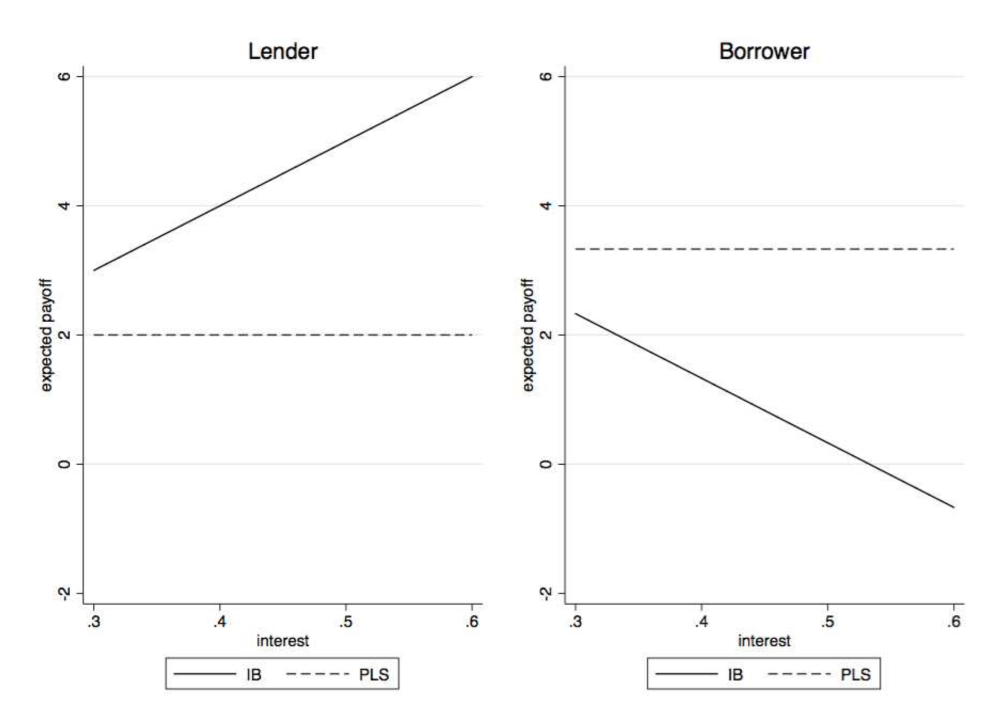

Figure 3.1: Expected Payoffs of Risk-Neutral Lenders and Borrowers. Source: Author's calculation.

We used a bi-directional Becker-DeGroot-Marschak (BDM, henceforth) procedure for lenders' decisions in the game. In the beginning of each round, the lender was informed the range of interest rates and that the true interest rate is within that range. Although we focus our investigation on the 30-60 percent range, we used both the 0-30 percent and the 30-60 percent ranges for four periods each, alternating between ranges. As illustrated in Figure 3.2, lenders made two decisions in each round. First, they chose a cutoff interest rate from the range. For example, if the range was 30-60 percent, a lender might choose a cutoff of 40 percent. Second, the lenders chose whether to invest the 10 points using IB or PLS for all interest rates equal to or above the cutoff interest rate. The lenders might choose IB for all interest rates equal to or above 40 percent and thus PLS for all interest rate below 40 percent. We label such choice as a *switching to IB*. The lenders might also choose PLS for all interest rates equal to or above 40 percent. We label such choice as a *switching to PLS*.

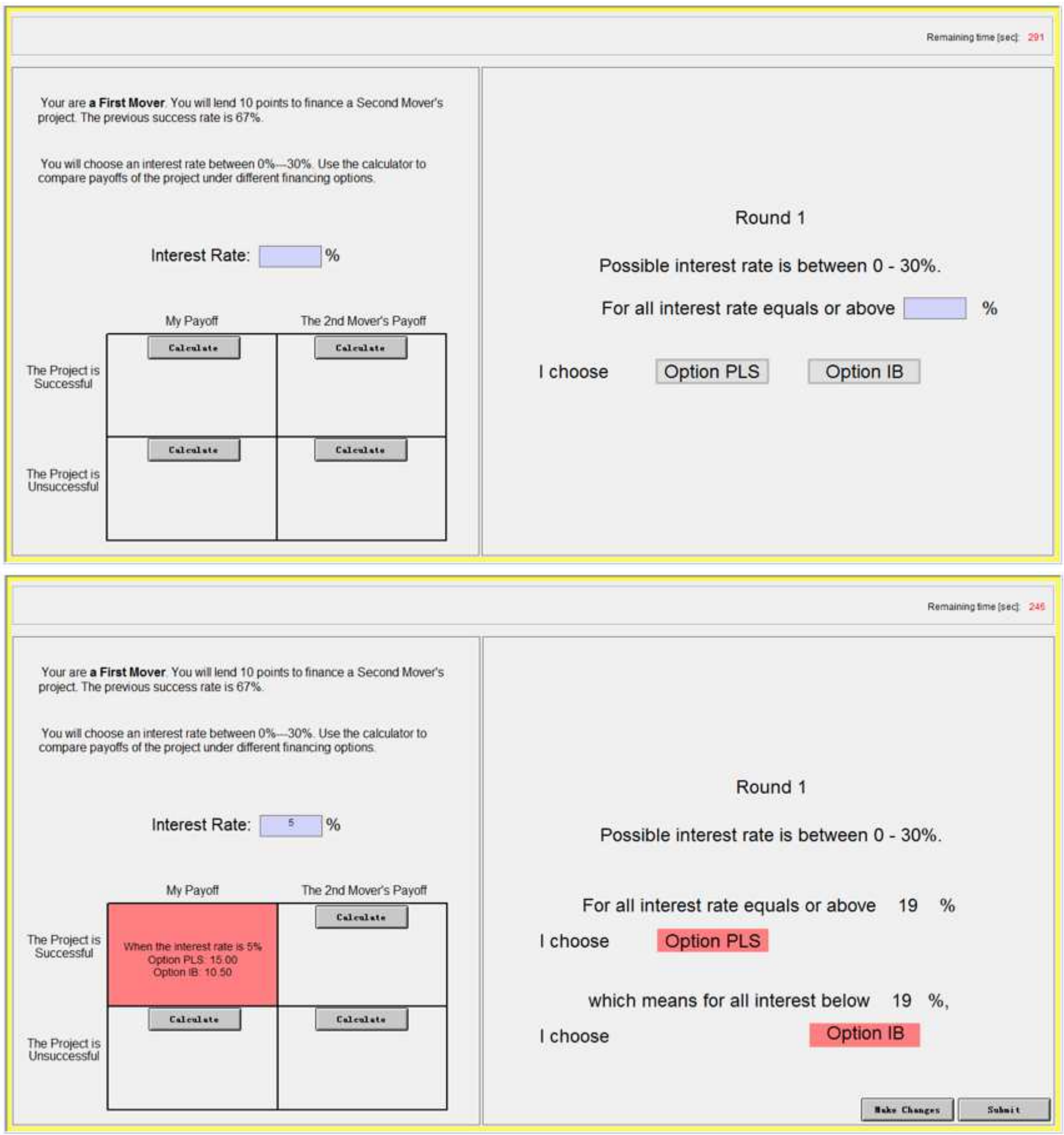

Figure 3.2: User Interface for Lenders in the Baseline Sessions

It is important to note that a lender could choose any interest rate, including the end-points, as a cutoff interest rate in the specified range of interest rates. If a lender switched to IB at a cutoff interest rate of 30 percent, the lender chose to invest in IB for all interest rates in the range. If a lender switched to IB at a cutoff interest rate of 60 percent, we assume the lender chose to invest in PLS for all interest rates in the 30-60 percent range. Similarly, we assume a lender chose to invest in IB for all interest rates in the 30-60 percent range if the lender switched to PLS at a cutoff interest rate of 60 percent.

We implement the bi-directional BDM for two main reasons. First, this design provides the lenders some degree of freedom to choose the financing arrangements throughout the range of interest rates. For example, the lenders can choose IB or PLS for all interest rates in the range, or the lenders can choose IB (PLS) for some interest rates before switching to PLS (IB). Second, this design accommodates different behavioral implications. A risk-averse or a risk-neutral payoff-maximizing lender will choose IB for all interest rates in the 30-60 percent. A prosocial lender might choose to invest in PLS as the interest cost imposed on her borrowers is quite large in this range. We discuss in detail two prosocial motives in Section 3.

Recall that lenders make financing decisions by either switching to IB or switching to PLS at some cutoff interest rates for 4 rounds. The implication of this design is that there is a possibility that lenders' financing choices differ across rounds. Specifically, the direction of switching may differ across rounds. For example, a lender might choose to switch to IB at a particular cutoff interest rate in the first round and then the lender might choose to switch to PLS in the following round. We identify lenders who made consistent choices across rounds, particularly in the last two or three rounds. A lender was consistent in n rounds if the switching directions were similar across the n rounds. We find that about 71 percent of lenders made consistent choices for at least 3 rounds and about 80 percent of lenders made consistent choice for at least 2 rounds.

We implement a between-subject design with a baseline treatment and a religious message treatment. In the religious message treatment, the subjects were shown a Quranic text about the prohibition of interest in lending: *"And whatever you lay out as usury, so that it may increase in the property of men, it shall not increase with Allah, and whatever you give in charity, desiring Allah's pleasure – it is these (persons) that shall get manifold,"* quoted from Surat Ar-Rum verse 39. Figure 3.3 illustrates the user interface shown to lenders in the message treatment.

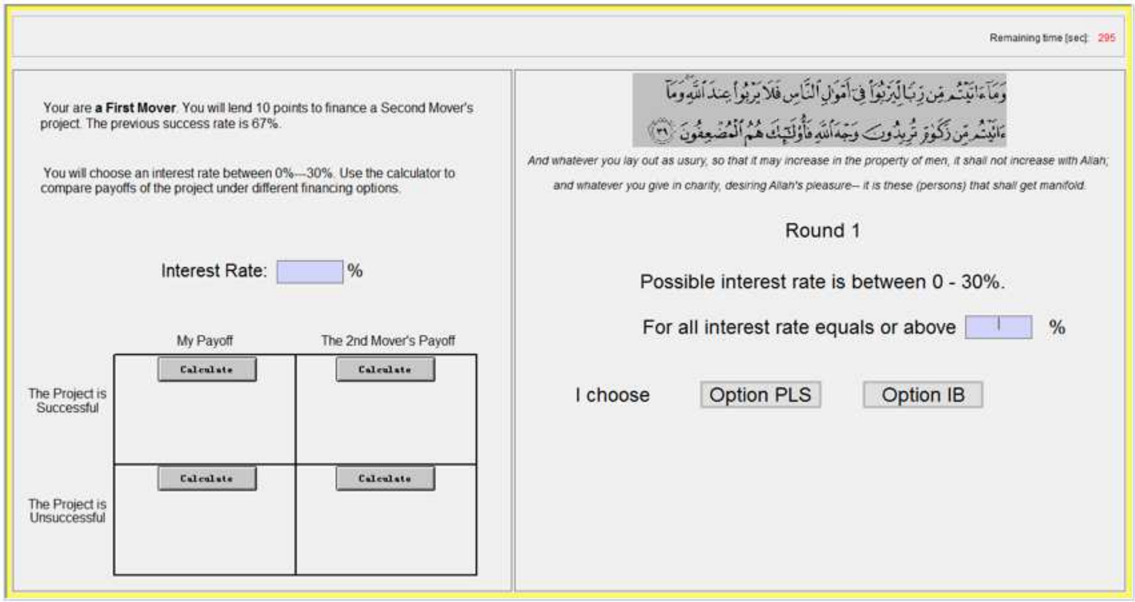

Figure 3.3: User Interface for Lenders in the Religious Message Sessions

As shown in Figure 3.2 and Figure 3.3, we equip subjects with a "mouse-lab" calculator to calculate the lender's and the borrower's payoffs under different success scenarios. To use the calculator, a lender entered a specific interest rate and clicked one specific cell at a time to reveal the payoffs. The two columns in the first row contains information about the lender's and the borrower's payoffs when the project is successful, while the columns in the second row contains information about pay offs when the project is unsuccessful. Each cell reveals the payoffs under IB and PLS arrangement.

In each session one round was randomly chosen for payment after the lenders made decisions for 8 rounds. Given the range of possible interest rates in the chosen round, a market interest rate was randomly drawn. The lender's decision in the chosen round and the market interest rate determined the financing arrangement to be implemented. For example, suppose that the chosen round was round 3, the chosen market interest rate is 50 percent, and a lender switched to IB at a cutoff interest rate of 40 percent. Then, IB would be implemented because the market interest rate is higher than the lender's cutoff interest rate.

Suppose that another lender in the session switched to PLS at a cutoff interest rate of 40 percent in the chosen round, then PLS would be implemented for that lender.

The randomly chosen round and the corresponding interest rate were written on a blackboard, and this written information were covered with dark-colored papers. These numbers were revealed after all lenders finished making choices in round 8. Each borrower discovered the financing arrangement that her matched lender has chosen for her given the randomly chosen round and interest rate. Each borrower then completed a project. Both lenders and borrowers know the outcome of the project, and thus their payoffs, only at the end of the experiment.

The subjects participated in a training round before each subject was assigned a role. During the training round, the subjects answered 4 hypothetical questions and learned to use the calculator. Specifically, subjects answered two hypothetical questions about the lender's and the borrower's payoff in the investment game. Subjects also answered two hypothetical questions about how a lender's choice and a randomly chosen market interest rate determines the lender's and the borrower's payoffs.

We conducted a post-experiment survey to obtain quantitative measures of individuals' characteristics, such as intrinsic religiosity, attitudes toward risky activities, family income, and sociodemographic information. We used questions from the DOSPERT scale developed by Blais and Weber (2006) to obtain a measure of individuals' risk attitudes. We used the religiosity of Islam scale (RoIS) developed by Jana-Masri and Priester (2007) to obtain a measure of individuals' religiosity. This Qur'an based survey include questions about beliefs and practices, which are important to measure religiosity of Muslims. We also measure religiosity using the centrality of religiosity scale (CRS-5) developed by Huber and Huber (2012).

The timing of the post-experiment survey, particularly the religious survey, is purely a design choice. A pre-experiment survey may introduce framing that may affect the subjects' decisions in the main experiment. On the other hand, we acknowledge that the religious message may affect the lenders' responses in the post-experiment survey. We find evidence that, as we discuss in Section 4, survey responses are quite balanced across treatments.

3.3. BEHAVIORAL HYPOTHESIS

In the previous section we show that the lender's expected payoff from IB is greater than her expected payoff from PLS for all interest rates in the 30-60 percent range. Moreover, the lender is always guaranteed a positive return from choosing IB. Thus, a payoff-maximizing lender would always choose IB for all interest rates in the 30-60 percent range. Choosing PLS is therefore the prosocial option. In this section we discuss two ways that a lender can be prosocial and our behavioral hypothesis.

First, a lender can appear prosocial by choosing PLS when the costs to them in terms of forgone payoff is low. For example is a lender who switches to IB at a cutoff interest rate of 45 percent. The lender protects her borrower from a loss but she does so when the forgone payoff is relatively lower. This lender chooses IB that guarantees herself a certain payoff when the cost is high. We label such lenders who are prosocial when costs to them are low as *protecting myself.* On the other hand, lenders can be prosocial when the costs are high. For example is a lender who switches to PLS at a cutoff interest rate of 40 percent. This lender protects the borrower from a loss even though her forgone payoff is relatively large. We label lenders who are prosocial when the costs are high as *protecting others.* We hypothesize that the religious message makes lenders more prosocial by protecting their borrowers from losses. In other words, we expect a lower proportion of lenders who switches to IB in the message treatment.

Hypothesis 1: there is a smaller proportion of lenders who switched to IB in the message treatment.

Recall that a lender selects the financing arrangement by either switching to IB or to PLS and by choosing a cutoff interest rate. We now discuss our hypothesis regarding the effects of the message on the choice of the cutoff interest rate. We first discuss protecting-other lenders who switch to PLS at particular cutoff interest rates. Figure 3.4 illustrates a lender who chooses to switch to PLS for all interest rates equal to or above 42 percent. The solid red line depicts the lender's expected payoffs for each interest rate. The lender forgoes sure payoffs for all interest rates equal to or above 42 percent for to protect her borrower. Choosing a lower cutoff interest rate is a more altruistic action as the lender protects the borrower for a larger interval. Thus, we hypothesize that the average cutoff interest rate is lower in the message treatment:

Hypothesis 2: the cutoff interest rate for those who switched to PLS is lower in the message treatment.

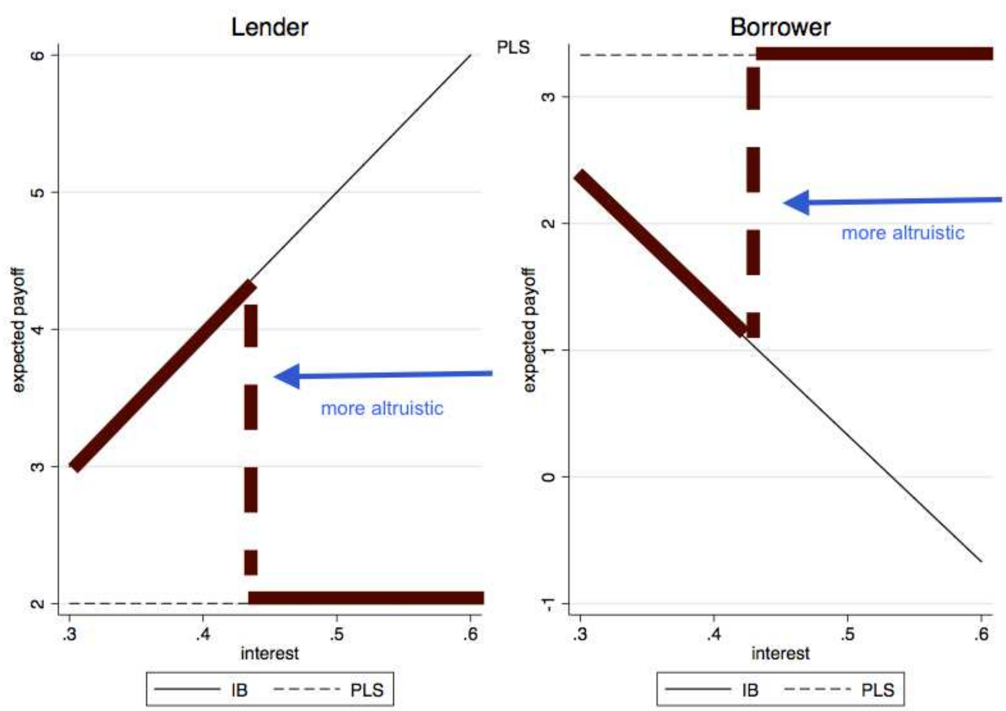

Figure 3.4: Lender's and Borrower's Expected Payoffs from Switching to PLS. Source: Author's calculation.

Hypothesis 2 is counterintuitive to the established theory of altruism as it shows that lenders help their borrowers when the cost imposed to himself is high. Andreoni and Miller

(2002) shows that subjects are rational altruists as subjects are more altruistic when the cost is relatively low. The main difference between our setting and theirs is that subjects' counterparts do not face any risk that could harm them. We conjecture that knowing that the counterpart is facing a risk that may harm her can alter the decision maker's preference for altruism despite a relatively high cost of being altruistic.

Lenders can also switch to IB at a cutoff interest rate. Figure 3.5 illustrates a lender who switches to IB for all interest rates equal to or above 46 percent. We can observe that the lender secures a certain and higher expected payoff when the forgone payoffs from choosing PLS are high. A more altruistic choice for such lender is to choose a higher cutoff interest rate. In other words, the lender covers a larger interest rate interval before protecting herself. Thus, we hypothesize that average cutoff interest rate for those who switch to IB is higher in the message treatment:

Hypothesis 3: the cutoff interest rate for those who switched to IB is higher in the message treatment.

We construct a variable that summarizes a lender's switching decision and choice of cutoff interest rate. This variable, which we refer to as the lender's expected gain, is the sum of the lender's expected payoff for each interest rate in the range. Thus, the expected gain is a function of the switching decision and the chosen cutoff interest rate. Specifically, the expected gain of a lender who switches to PLS at a cutoff interest rate $\overline{r}$ is then given by:

$$E\left[Y_{(l,u)}\right] = \frac{1}{31}\left[\sum_{r=30}^{\overline{r}-1} E\left[IB\left(r\right)\right] + \sum_{r=\overline{r}}^{60} E\left[PLS\left(r\right)\right]\right]. \tag{3.1}$$

while the expected gain of a lender who switches to IB at a cutoff interest rate $\overline{r}$ is given by:

$$E\left[Y_{(l,u)}\right] = \frac{1}{31}\left[\sum_{r=30}^{\overline{r}-1} E\left[PLS\left(r\right)\right] + \sum_{r=\overline{r}}^{60} E\left[IB\left(r\right)\right]\right] \tag{3.2}$$

The expected gain variable has two important properties. First, a lender's expected gain

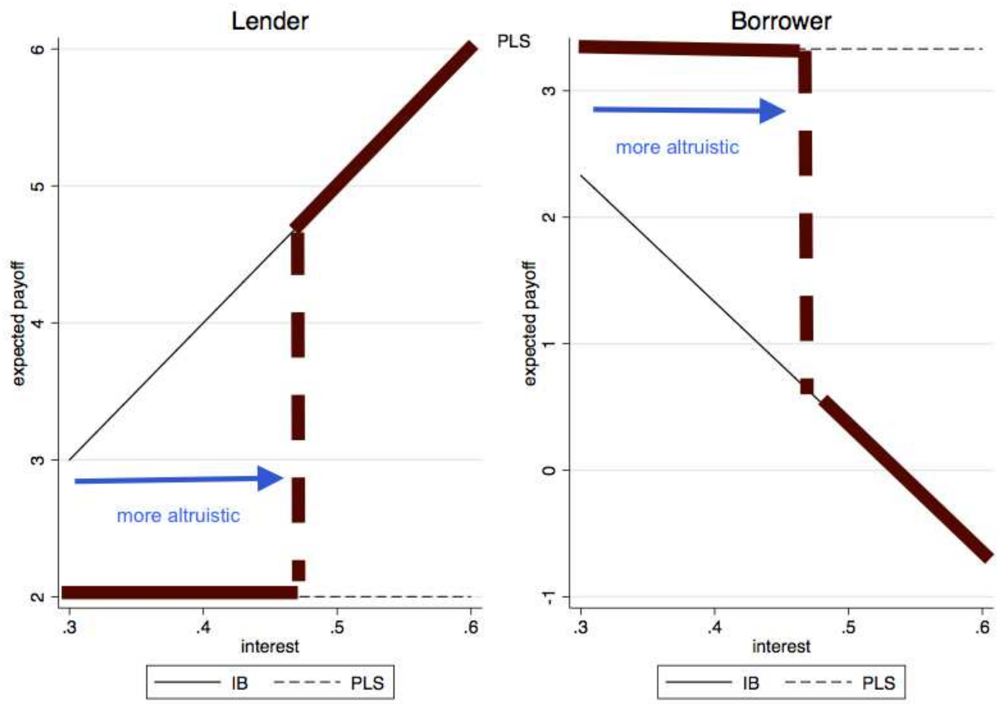

Figure 3.5: Lender's and Borrower's Expected Payoffs from Switching to IB. Source: Author's calculation.

is the inverse of the borrower's expected gain, providing a straightforward comparison. A higher lender's expected gain implies a lower borrower's expected payoff. More importantly, the variable highlights the tension that the lender faces when making a decision: a higher gain for himself on the expense of the borrower's gain, vice versa. Consider the graphs in the first column of Figure 3.6. A lender guarantees himself the highest expected gain by switching to IB for interest rates equal to and above 30 percent. This choice, however, results in borrower's lowest expected gain as shown on the second column.

Second, the value of the expected gain increases with the number of selfish financing arrangements chosen. For example, consider a lender who switches to IB at a cutoff interest rate of 45 percent and another lender who switches to IB at a cutoff interest rate of 30 percent. The graph in the first column of Figure 3.6 shows that the expected gain for the second lender is higher than the one for the first lender. This property also shows that the

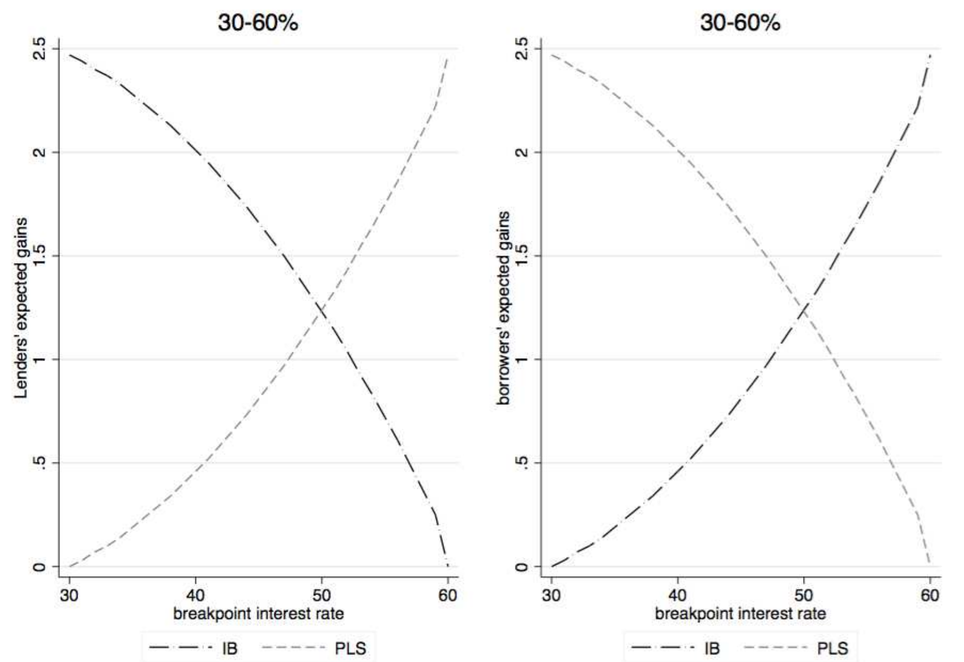

Figure 3.6: Lender's and Borrower's Expected Gains Based on the Lender's Choice. Source: Author's calculation. Notes: the dashed black line (dashed grey line) in the first column indicates the lender's expected payoff from choosing IB (PLS) for all interest rate equal or higher than breakpoint interest rate.

expected gain of a lender who switches to IB at relatively high cutoff interest rate is lower than the expected gain of a lender who switch to PLS at the same cutoff interest rate. As shown in Figure 3.6, the expected gain from switching to IB is decreasing with interest rate while the expected gain from switching to PLS is increasing with interest rate.

Hypotheses 2 and 3 imply a straightforward prediction on the lenders' expected gains. Specifically, the lenders' expected gain is lower in the message treatment. Hypothesis 2 suggests that lenders who switch to PLS choose a lower cutoff interest rate. As shown in Figure 3.6, this implies a lower expected gain. Hypothesis 3 suggests that lenders choose a higher cutoff interest rate before switching to IB, which also implies a lower expected gain. Hypothesis 1 also suggests a lower lenders' expected gain in the message treatment owing to a higher proportion of lenders who switch to PLS.

Hypothesis 4: the lenders in the message treatment earned lower expected

gains.

3.4. RESULTS

We conducted our experiments at three different locations: 1) The Faculty of Economics and Business Computer Lab, Universitas Gadjah Mada, Yogyakarta, Indonesia; 2) Ningxia University, Ningxia Hui Autonomous Region, China; and 3) Vernon Smith Experimental Laboratory, Shanghai Jiao Tong University, Shanghai, China. We ran 6 baseline and 6 treatment sessions in Indonesia and China and we ran 4 baseline and 4 treatment sessions in Shanghai. In each session, there are on average 10 lenders and borrowers pairs.[5]

The principal treatments were programmed using z-Tree (Fischbacher, 2007). In each session subjects participate and make decision anonymously. In each country the same experimenter read the instructions in the beginning of each session to ensure that everyone had common knowledge of the decision tasks. No form of communication is possible during the experiment. On average, one session lasted for 90 minutes. Subjects in Indonesia receives a Rp25,000 show-up payment and, on average, earn Rp112,500. Subjects in China receives a RMB10 show-up payment and, on average, earn RMB65.

Table 3.1 reports the summary statistics for the baseline and the treatment session. The majority of the lenders were female, which accounts for about 65 percent of the total lenders. The subjects were about 20 years old and they have completed on average 2.7 years of college. About 1 in every 4 lenders worked full or part time. Overall, the descriptive statistics suggest that the demographic characteristics in the baseline and the treatment sessions are quite balanced. We conducted a survey at the end of each experiment. The survey included questions to obtain measures of intrinsic religiosity and attitudes toward risky activities. One may be concern that the levels of religiosity were higher in the message treatment,

[5]We ran a couple of session in Indonesia with less than 10 lender-borrower pairs, therefore we ran other sessions with 11 or 12 lender and borrower pairs. In China, We ran a session with 7 lender-borrower pairs, therefore we ran an extra session with 3 lender-borrower pairs.

since lenders might chose answers that justified their choices during the experiment. Table 3.1 show no evidence that the means of the religiosity were significantly different across treatments.

Table 3.1: Summary Statistics of the Muslim Lenders

Variables	Baseline		Message	
	Mean	Std. Deviation	Mean	Std. Deviation
Gender, 1 if Male	0.38	0.49	0.33	0.47
Age	20.35	1.84	20.91**	1.91
College year completed	2.72	1.88	2.99	1.61
Work Full or Part Time	0.26	0.44	0.33	0.54
Risk Attitude scale	56.18	12.16	54.46	12.93
Risk Attitude scale, standardized	0.03	1.03	-0.03	0.97
Religiosity scale	0.65	0.16	0.67	0.16
Religiosity scale, standardized	-0.06	1.02	0.06	0.98
Number of lenders	120		123	

Source: authors' calculations of the experimental data.
Notes: risk attitude and religiosity scales are standardized with mean 0 and variance 1 in each country. The signs *, **, *** indicates significance of t-test at 10, 5, and 1 percent.

We test the first hypothesis that the religious message led to a lower proportion of lenders who switched to IB. We regress the proportion of individuals who switched to IB on a message dummy and two location dummies. We cluster the standard errors at the individual level since each lender made a total of 4 decisions. Table 3.2 summarizes the effects of religious message on the choice of investment arrangement. In Column A3 of Table 3.2, we can observe that the the proportion of lenders who switched to IB is significantly lower by about 8 percentage points in the message treatment. We obtain similar result if we restrict the sample to decisions in the last round.

Our data do not support hypotheses 2 and 3 that the religious message affects the cutoff interest rate. As shown in Column B3 and Column C3 of Table 3.2 the differences in the means of cutoff interest rate are not significantly different to zero. However, our data provide a support of hypothesis 4 that the lenders' expected gain is lower in the message treatment. The difference in the lenders' expected gains across the two treatments is about 0.132 point as shown in Column A3 of Table 3.2. The lower expected gains in the message treatment

Table 3.2: The Effects of the Religious Message on Choices and Expected Gains in Indonesia and Ningxia: Range 30-60 Percent

	A: All Decisions			B: Switching to IB			C: Switching to PLS		
	1: Baseline	2: Message	3: Difference	1: Baseline	2: Message	3: Effect	1: Baseline	2: Message	3: Effect
Switch to IB	0.644	0.563	-0.081**						
Expected gain	1.308	1.176	-0.132**	1.688	1.647	-0.041	0.623	0.569	-0.053
Breakpoint interest rate				43.194	43.732	0.538	40.778	39.679	-1.098
Total observations	480	492	972	309	277	277	171	215	386
Last Period									
Switch to IB	0.667	0.528	-0.138**						
Expected gain	1.394	1.206	-0.188**	1.756	1.773	0.016	0.669	0.570	-0.098
Breakpoint interest rate				42.112	41.738	-0.374	41.225	39.465	-1.759
Total observations	120	123	333	80	65	145	40	58	98

Source: authors' calculations of the experimental data.
Notes: the critical values for a one-tailed test are: 1.290 (10 percent), 1.660 (5 percent), and 2.364 (1 percent). The signs *, *, *** indicates significance at 10, 5, and 1 percent for one-tailed tests.

is mainly driven by a lower proportion of lenders who switched to IB. The results thus far show that the message motivated subjects to be more prosocial and, in particular, to protect the borrowers from losses even though the costs are high.

We now discuss a model specification to investigate the effects of the religious message on the lenders' expected gains. Our main specification is:

$$Y_{is} = \beta Message_i + \Gamma X_i + \epsilon_{is} \tag{3.3}$$

where Y_{is} indicates lender i's expected gains in round s, $Message$ is an indicator for the message treatment, and the vector X includes three dummies for the rounds of play, a dummy for gender, an indicator for experiments in China, years in college, age, and religiosity. There is an established finding in the literature that religiosity and measures of risk attitude are negatively correlated (Miller and Hoffmann, 1995; Noussair et al., 2013). Indeed, our data suggest significant and negative correlation between religiosity and risk attitude scale. The estimated religiosity parameter may be biased downward if we do not control for risk attitude scale. Thus, we include a measure of risk-attitude as one of the control variables in the estimation.

For ease of interpretation, we standardized the religiosity and the risk attitude scales by country to have means of zero and a standard deviation of 1. We also estimate a specification in which we allow the effects of the message to vary by religiosity. Specifically, we include an interaction between the message indicator and the standardized religiosity measure. We cluster the standard errors at the individual level in all of our specification. Table 3.3 reports the results of the regression analysis.

In Column 1 of Table 3.3 we can observe that the lenders' expected gains are on average lower in the message treatment, while in Column 2 of Table 3.3 we can observe a negative relationship between expected gains and the lenders' religiosity. We also estimate a specification that includes an interaction between the message indicator and lenders' religiosity

Table 3.3: The Effects of Religious Message on Expected Gains: Range 30-60 Percent

	1: Model 1	2: Model 2	3: Model 3	4: Model 4	5: Model 5
Message	-0.132*	-0.121*	-0.125*	-0.092	-0.105
	(0.071)	(0.070)	(0.069)	(0.069)	(0.067)
Religiosity		-0.071**	0.015	0.029	0.014
		(0.034)	(0.043)	(0.042)	(0.041)
Message · Religiosity			-0.170***	-0.180***	-0.170***
			(0.064)	(0.064)	(0.065)
Observations	972	972	972	972	972
Controls	N	N	N	Y	Y
Country FE	N	N	N	N	Y

Source: authors' calculations of the experimental data.
Notes: standard errors are clustered at individual level. The signs *, *, ** indicates significance at 10, 5, and 1 percent. Other regression covariates not shown in the table are dummies for rounds, a dummy for gender, a dummy for China, years in college, age, and standardized religiosity score.

and we report the estimated parameters in Column 3 of Table 3.3. We conjecture that the more religious lenders might be more responsive to the Islamic prohibition towards interest. Indeed, the result in Column 3 of Table 3.3 shows that the effect of the message is stronger among the more religious lenders. In Column 4 and Column 5 of Table 3.3 we augment the specification with control covariates, such as the standardized risk-attitude scale, age, gender, and years in college, and country fixed effects. The results show that the estimated parameters are not sensitive to the inclusion of the covariates.

We depict the marginal effects of the religious message on lenders' expected gains by the level of religiosity in Figure 3.7. Recall that the religiosity scale is normalized within each country to have a mean of 0 and standard deviation of 1. Thus, the figure shows that the effects of religious message are significant among lenders with religiosity level above the mean in each country.

We perform regressions using subsamples to investigate whether the results are sensitive to changes in the regression sample. First, we perform regression analysis for lenders who made consistent choices across rounds. We find that about 71 percent of the lenders was

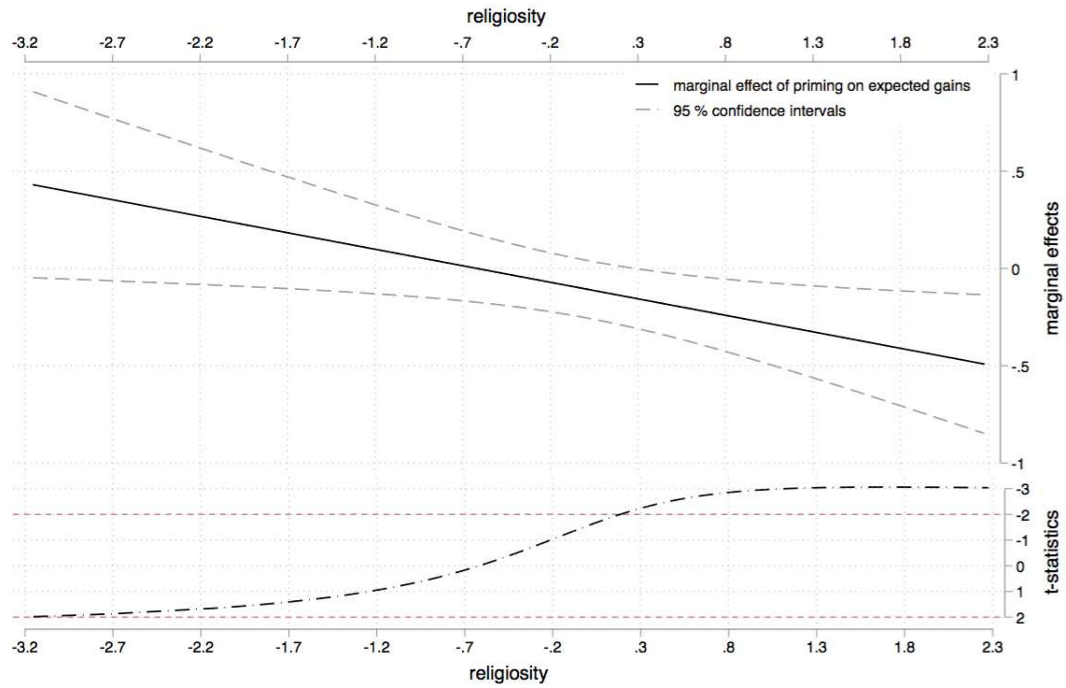

Figure 3.7: The Marginal Effects of Religious Message on Lenders' Expected Gains. Source: author's calculation of the experimental data.

consistent in at least three rounds and about 80 percent of the lenders was consistent in at least two rounds. The results in C.2 shows that the results are robust when we restrict the sample to include only lenders who were consistent. Second, we perform regression analysis to investigate the effect of the religious message in each round. Table C.1 shows that the effect of religious message is stronger as the round progresses. The estimated parameter for the interaction term between religious message and religiosity is significant starting in rounds 3 & 4 and the absolute magnitude of the treatment parameter increases in the later rounds.

3.4.1 Calculator Usage

We now turn to discuss the lenders' calculator usages to understand the lenders' thought process. Recall that each lender could use the calculator to calculate and to reveal payoffs in different scenario. For any interest rate in the range, a lender could calculate her payoffs or her borrower's payoffs if the project is successful or unsuccessful. Table 3.4 reports the summary statistics of the calculator usages. The number in a specific cell indicates, on average, the number of times the lenders revealed payoffs for a particular scenario.[6] The results show that the lenders cared more about their own payoffs. On average, the lenders revealed her payoffs twice as much as their borrower's payoffs in all scenario. We test whether the religious message affected calculator usages and we find no significant effect of the religious message. This result indicates that the lenders in the message treatment used the calculator as much as the lenders in the baseline treatment. This is an indication that the lenders who observed the Quranic quote also made informed choices.

Table 3.4: Summary Statistics of Calculator Usage by the Lenders

	Own: Successful	Own: Unsuccessful	Other: Successful	Other: Unsuccessful	Observations
Baseline	5.017	4.063	2.146	1.529	480
Message	4.967	3.986	2.128	1.671	492

Source: authors' calculations of the experimental data.
Notes: we regress the calculator use on a message indicator and two location dummies to test the effects of the message on calculator usage. The standard errors are clustered at individual level. The signs *, *, ** indicates significance at 10, 5, and 1 percent for two-tailed tests.

To further understand the lenders' thought process, we now ask the following question: do calculator usages influence the choice to switch to IB and thus the expected gains in the baseline and the message treatment? To answer this question, we regress the lenders' choices to switch to IB and the lenders' expected gains on the number of calculator clicks for all four scenarios, an indicator for message treatment, its interaction with the number of calculator clicks for each scenario, and other covariates. Table 3.5 reports the results of the regression analysis.

We first discuss the results for the baseline treatment in Row 1 and Row 3. The results show that the calculator usages influenced the choice to switch to IB and the expected gains. In Column 1 and Column 4 we can observe that a higher number of clicks to reveal lenders' successful payoff or borrower's unsuccessful payoff is associated with a lower likelihood of

[6]For each lender, we record the number of times the lender reveals different scenario for every interest rate. We sum the number of times the lender reveals the payoffs for a particular scenario. We then take the average for a particular scenario across all lenders.

switching to IB. The intuition of the result in Column 1 is the following. The lenders could observe that the payoffs from IB and PLS do not differ that much if the project is successful. For some interest rates, the payoff from PLS is even higher than the payoff from IB. This might influence the lenders to choose PLS. The result in Column 4 suggests that the lenders' observation of the borrowers' payoff in an unsuccessful project motivated the lenders to choose a prosocial action. In Column 2, on the other hand, a higher number of clicks to reveal the lenders' unsuccessful payoff or the borrowers' successful payoff is associated with a higher likelihood of switching to IB. This shows that, as lenders were more aware of the risks, they were more likely to choose the option that guarantees a certain payoff. The correlation between calculator usage and expected gains are analogous.[7]

Table 3.5: Correlation between Calculator Usage and Choices: Range 30-60 percent

	Own Success	Own Unsuccess	Other Success	Other Unsuccess
Switch to IB, Baseline	-0.043**	0.052**	0.022	-0.046**
	(0.021)	(0.022)	(0.020)	(0.023)
Switch to IB, Priming	0.023	-0.002	-0.008	-0.040*
	(0.019)	(0.019)	(0.019)	(0.021)
Expected gains, Baseline	-0.081**	0.078**	0.048	-0.058
	(0.031)	(0.031)	(0.036)	(0.038)
Expected gains, Priming	0.050	-0.040	-0.012	-0.038
	(0.033)	(0.035)	(0.034)	(0.039)
Observations	972	972	972	972
Controls	Yes	Yes	Yes	Yes

Source: authors' calculations of the experimental data.
Notes: we use a linear probability model on the regression of the choice of IB on covariates. Standard errors are clustered at individual level. The signs *, **, *** indicates significance at 10, 5, and 1 percent. Other regression covariates not shown in the table are dummies for rounds, a dummy for gender, dummies for location, and a treatment dummy.

The results reported in Row 2 and Row 4 of Table 3.5 show that calculator usages are not correlated with the lenders' financing choices as well as the lenders' expected gains in the message treatment. This result seems to indicate that the higher proportion of prosocial choices in the message treatment is mainly influenced by the religious message. The takeaway from the analysis of the calculator usages is that the lenders made calculated and informed

[7]We perform similar analysis for each country. We find that the signs of the estimated parameters are consistent across country.

choices to take risks for the borrowers even after the lenders observed the religious message.

3.4.2 Cross-country differences

In our main experiment, we investigate the effect of religious message on the choice of financing. We find that the lenders in the message treatment were more prosocial as they took more risks to protect their borrowers from losses. We now analyze cross-country differences to investigate how the message affected choices in each country. First, we discuss the summary statistics of the lenders in each country which we report in Table C.3. We highlight several important variations across countries. In Indonesia, the means of the risk attitude measure is significantly lower in the message session, which can be explained by the relatively higher number of female lenders.[8] As expected, the mean of the religiosity measure in Indonesia is the highest.

We now test the hypotheses for each country. We regress the variable of interest with a treatment dummy and we cluster the standard errors at the individual level. Table 3.6 presents the results of the estimations. We find different patterns on the effects of the message on choices. In Indonesia we find evidence that supports our first hypothesis that there is a higher proportion of protecting-other lenders. The share of lenders who switched to IB is significantly lower by about 10 percent in the message treatment. However, a lower share of switching to IB in the message treatment does not correspond to a significantly lower expected gain. This seems to be driven by a weakly lower cutoff interest rate among lenders who switched to IB in the message treatment.

We observe a distinct effect of the religious message on choices in China. We find no support of hypothesis 1 as the proportion of lenders who switched to IB across treatments are not statistically different. However, in Row 2, we find evidence to support hypothesis 3. The message seems to affect the choice of cutoff interest rate among the protecting-myself lenders. Specifically, the lenders in the message treatment chose the risky PLS for a larger

[8] In a separate analysis, we find a significant negative correlation between risk attitude measure and being female.

Table 3.6: The Effects of Religious Message on Choice of IB in All Rounds: Range 30-60 percent

	Indonesia			Ningxia			Shanghai		
	Baseline	Message	Effect	Baseline	Message	Effect	Baseline	Message	Effect
Switch to IB	0.62	0.53	-0.10*	0.71	0.68	-0.03	0.92	0.91	-0.01
Expected payoffs	1.15	1.03	-0.12	1.46	1.32	-0.13*	1.98	2.05	0.06
Total observations	240	252	492	240	240	480	160	160	320
				Switching to IB					
Breakpoint interest rate	46.82	47.74	1.29	43.1	45.28	2.18**	37.35	36.82	-0.52
Expected gains	1.35	1.25	-0.10	1.70	1.53	-0.17**	2.03	2.06	0.03
Total observations	151	134	285	170	162	332	147	145	292
				Switching to PLS					
Breakpoint interest rate	42.97	42.33	-0.64	44.54	44.17	-0.36	50.38	54.67	4.28
Expected gains	0.81	0.78	-0.03	0.88	0.89	0.01	1.45	1.89	0.44
Total observations	89	118	207	70	78	148	13	15	28

Source: authors' calculations of the experimental data.
Notes: the critical values for a one-tailed test with 100 degree of freedom are: 1.290 (10 percent), 1.660 (5 percent), and 2.364 (1 percent). The signs *, *, *** indicates significance at 10, 5, and 1 percent for one-tailed tests.

interest rate interval before switching to IB. These choices led to significantly lower expected gains in the message treatment as the lenders gave up some ranges that offer them certain payoff.

3.4.3 Religion or morality?

A straightforward follow-up question about the findings so far is, does the message affects behavior through the religious or moral precept? This is a quite challenging inquiry since the practical answer would be through both, as religion and morality are essentially intertwined. Nevertheless, we formally test this inquiry to gain a better understanding of the effect of the religious message. Specifically, we replicate our experiment with 80 non-Muslim lenders in Shanghai. About 95 percent of the lenders in the Shanghai sessions do not have a religion, while the other lenders are either Catholics, Buddhists, or other undisclosed religion. If the message works through religion, we expect to see no significant effect of the message among the lenders in Shanghai.

In Table 3.6 we can compare the results from Muslim lenders in Ningxia and ones from non-Muslim lenders in Shanghai. We can observe that most of the Shanghai lenders were quite selfish in comparison to the Ningxia lenders. First, more than 90 percent of the Shanghai lenders chose to switch to IB even in the baseline treatment. Second, these lenders switched to IB, on average, at a quite low cutoff interest rate. We find no effect of the message on financing choice and expected gains. This finding provides a strong support that the message works through religion.

3.4.4 Religion or general altruism?

Our last robustness check is to investigate the effects of the message on financing choice for the 0-30 percent range. In Figure C.1, we graph the expected payoffs from IB and PLS for each interest rate in this range. In this range IB is, in expectation, more beneficial for the borrowers particularly when the interest rate lower than 20 percent. Therefore, the religious message is not in line with an action that benefits the borrowers, which in this case is choosing IB. Does the message work when there is an apparent conflict between what the message commands and the action that benefits the others?

We test the effect of the message on choices for the 0-30 percent interest rate range. As discussed in Section 3, lenders made financing decisions for the 0-30 percent range for four rounds. In each round the lenders chose a cutoff interest rate and the lenders chose whether they want to switch to IB or PLS for all interest rate equal or above the chosen cutoff interest rate.

Table 3.7 summarizes the results for the 0-30 percent. First, we investigate the effects of religious message on the proportion of lenders who switch to IB. In Column 3, we can observe that the difference in the proportion across rounds is essentially zero. Among the lenders who switched to IB, the average cutoff interest rate is about 14 percent, and it is not significantly different across treatment. These choices imply that there is no significant difference in expected gains across treatments. The results shows that the religious message does not affect financing choice in the 0-30 percent range. This suggests that a religious message can motivate prosocial behavior when what is prescribed by the message is in line with being prosocial.

3.5. CONCLUSION

In this paper, we contribute to the literature by exploring the interaction between religion and prosocial risk taking. We investigate this interaction in a context of the Islamic prohibition against lending with interest. Although prohibited, many Muslims lend using interest-based arrangement owing to the global use of interest. Islamic law prescribes PLS arrangements as the alternative arrangement, through which lenders share the risk of investment. However, the adoption of such arrangement is quite low in Muslim majority countries. We establish an experiment in which the subjects were divided into lenders and borrowers. The lenders chose among two available financing arrangements, which resemble the conventional IB arrangement and the PLS arrangement. By choosing IB, a lender guarantees a certain and a higher payoff to the lenders. On the other hand, a lender shares the risk of the project with her borrower if the lender chooses PLS. To analyze the effect of religion on financing choices, we make salient the Islamic prohibition against the conventional IB arrangement in the message treatment.

Our analysis shows that in the message treatment there were more lenders who protected their borrowers even though the costs in terms of forgone payoffs were high. Specifically, a higher proportion of lenders switched to PLS at high interest rates in the message treatment. By switching to PLS at high interest rates, the lenders' forgo certain payoffs for high interest rates and thus earned lower expected gains in the message treatment. We find that the religious message effects vary with religiosity: the effects of the religious message were stronger and significant among lenders with religiosity above the mean level. Our cross-country analyses show that the effects of the message on taking up risks for the borrowers are quite consistent across the country. Finally, our robustness checks show two important insights about the effects of the message: the message works through the religion channel,

Table 3.7: The Effects of Religious Message on Choice of IB in All Rounds: Range 0-30 Percent

	A: All Decisions			B: Switching to IB			C: Switching to PLS		
	1: Baseline	2: Message	3: Effect	1: Baseline	2: Message	3: Effect	1: Baseline	2: Message	3: Effect
Switch to IB	0.517	0.510	-0.006						
Expected gains	0.460	0.448	-0.012	1.688	1.647	-0.041	0.702	0.671	-0.031
Breakpoint interest rate				14.488	13.345	-0.742	13.013	12.568	-0.044
Total observations	480	492	972	248	251	499	232	241	473

Source: authors' calculations of the experimental data.
Notes: the critical values for a one-tailed test with 100 degree of freedom are: 1.290 (10 percent), 1.660 (5 percent), and 2.364 (1 percent). The signs *, *, *** indicates significance at 10, 5, and 1 percent for one-tailed tests.

but not morality, and the message works when the choice prescribed by the message is the altruistic choice.

We close by highlighting several implications of our findings. First, our results provide a support to other experimental studies that find a positive relationship between religion and prosociality. We also contribute to the literature by investigating the effects of religious message on taking a risky prosocial action. Second, studies on giving find a negative relationship between giving and the costs of giving. We find evidence that religious message induce a higher proportion of subjects that help others even when the cost imposed to themselves is high. A more careful analysis about the rationality of taking up risks to help others is a topic of future research agenda. Third, our findings serve as evidence for IBF institutions that PLS is an acceptable investment arrangement even when the opportunity cost of choosing PLS is high. However, we believe that future works is necessary to address problems inherent in PLS arrangement, such as agency problems. Nevertheless, PLS is an alternative investment arrangement that microfinance institutions can offer. These institutions give lending to the poor, who face relatively high lending interest rates in the informal market. For example, a policy brief by Kneiding and Rosenberg (2008) shows that the global average of microcredit interest rates is 35 percent.

APPENDIX A

ADDITIONAL MATERIALS FOR FOR CHAPTER 1

Table A.1: Key Variables from the NLSY79

Variable	**Variable Title in NLSY79**	**Question Name**
father's education	highest grade completed by R's father	HGC-FATHER
mother's education	highest grade completed by R's mother	HGC-MOTHER
AFQT test score	AFQT percentile score, revised 2006	AFQT-3
years of schooling	highest grade completed as of May 1 (revised)	HGCREV79-10
hours worked	number of hours worked in past calendar year	HRSWK-PCY
weeks worked	number of weeks worked in past calendar key	WKSWK-PCY
income	total income from wages and salary in past calender year	Q13-5, Q13-5_TRUNC, Q13-5_TRUNC_REVISED
msa	is R's current residence in SMSA?	SMSARES
total enrollment	total enrollment in R's school in 10-1-79 - school survey	SCHSUR-17
number of teachers	number of full-time equivalent teachers in school - school survey	SCHSUR-43
teacher with Master's	percentage of full-time teachers with Master's/Doctor's - school survey	SCHSUR-44
teacher salary	annual salary beginning certified teacher with BA - school survey	SCHSUR-46

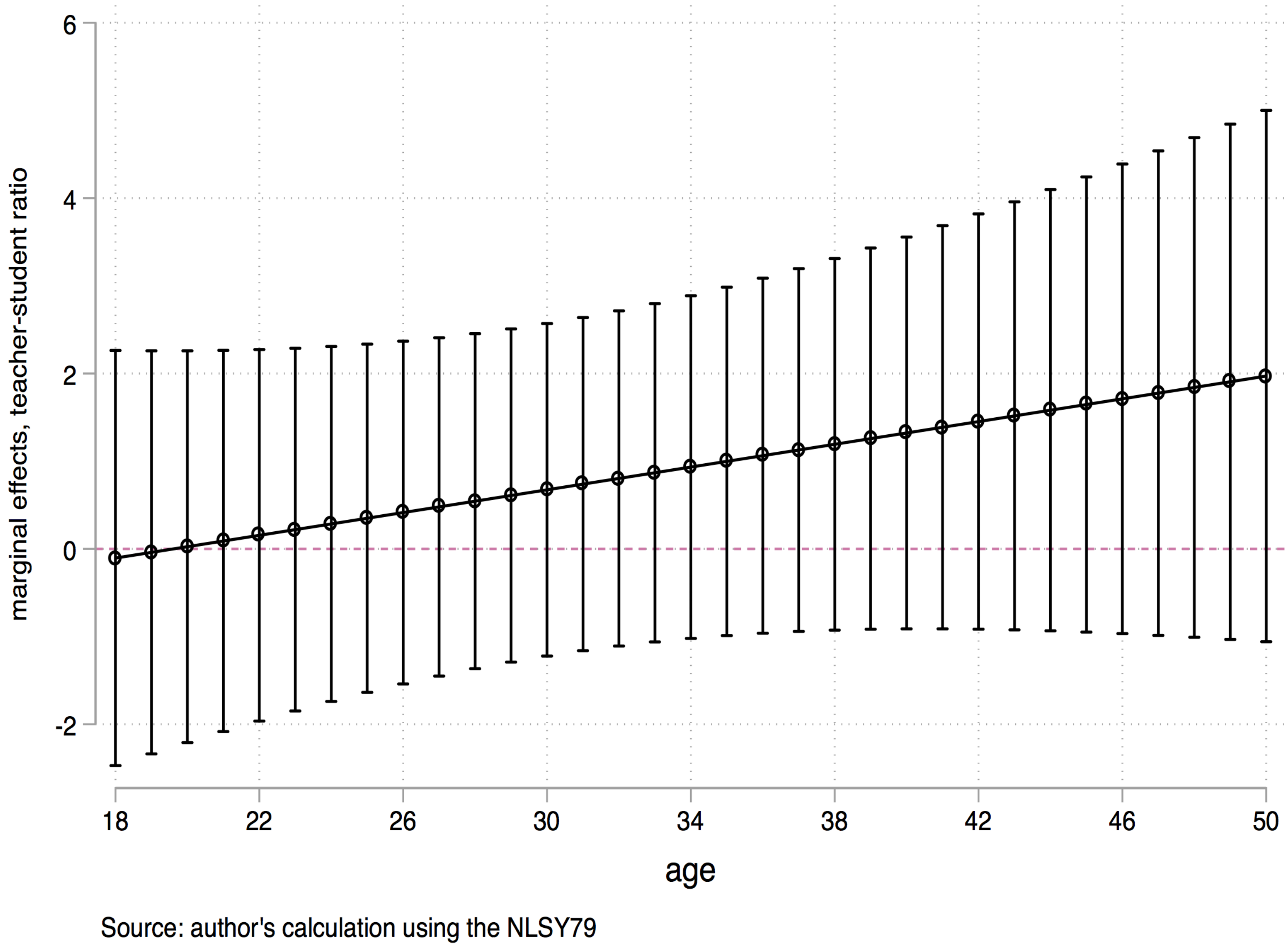

Figure A.1: Marginal Effects of Teacher-Student Ratio by Age. Source: author's calculation using the NLYS79. Notes: standard errors are clustered at the individual's level. The covariates used in the estimation model are years of schooling, quartic function of experience, a dummy for marital status, a dummy for residence in SMSA, census region dummies, interaction terms between education and the census region dummies, and time dummies.

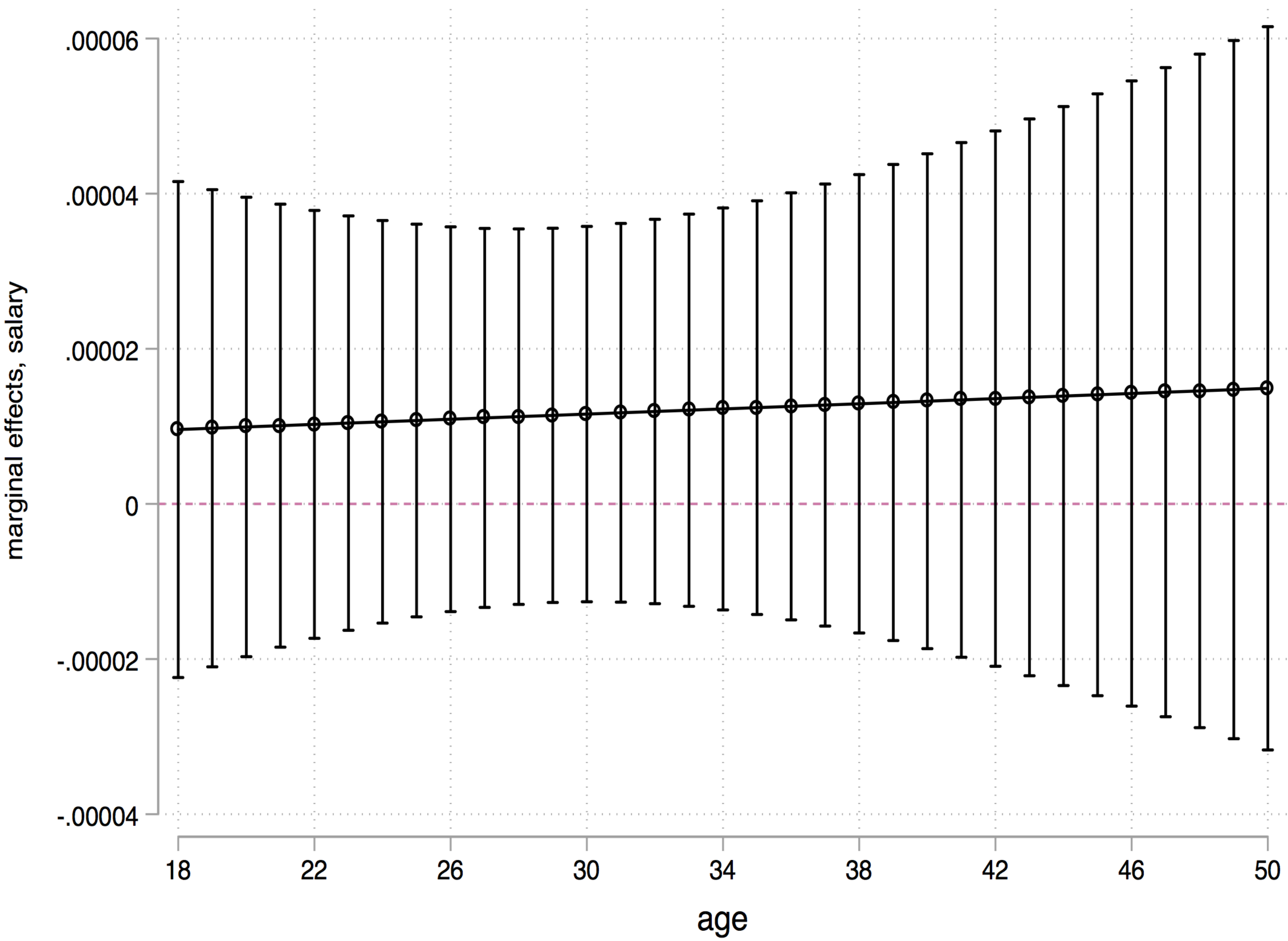

Figure A.2: Marginal Effects of Teacher Salary by Age. Source: author's calculation using the NLYS79. Notes: standard errors are clustered at the individual's level. The covariates used in the estimation model are years of schooling, quartic function of experience, a dummy for marital status, a dummy for residence in SMSA, census region dummies, interaction terms between education and the census region dummies, and time dummies.

APPENDIX B

ADDITIONAL MATERIALS FOR CHAPTER 2

Table B.1: Impacts of Scholar Selection on College-Going Outcomes: Fuzzy RD Estimates

	2009-2012 Cohorts						2009-2010 Cohorts					
Outcomes	μ	Full Sample	Intermediate (±100)	Narrow (±40)	Optimal Bandwidth	Range of Bandwidths	μ	Full Sample	Intermediate (±100)	Narrow (±40)	Optimal Bandwidth	Range of Bandwidths
1: Intermediate enrollment	0.854	0.016	0.023	0.026	0.016	44	0.852	0.032	0.039	0.069	0.107*	36
		(0.023)	(0.026)	(0.041)	(0.039)			(0.031)	(0.036)	(0.054)	(0.059)	
		[3,019]	[2,585]	[1,372]	[1,482]			[1,454]	[1,245]	[668]	[602]	
2: 2^{nd} year persistence rate	0.753	0.024	0.040	0.036	0.028	44	0.753	0.035	0.056	0.078	0.112	38
		(0.026)	(0.031)	(0.049)	(0.047)			(0.036)	(0.043)	(0.064)	(0.069)	
		[3,019]	[2,585]	[1,372]	[1,482]			[1,454]	[1,245]	[668]	[631]	
3: 3^{rd} year persistence rate	0.677	0.048*	0.064*	0.092*	0.126**	36	0.661	0.068*	0.093**	0.132*	0.202***	36
		(0.028)	(0.034)	(0.053)	(0.056)			(0.039)	(0.046)	(0.071)	(0.076)	
		[3,019]	[2,585]	[1,372]	[1,240]			[1,454]	[1,245]	[668]	[602]	
4: BA attainment, in 4 years	0.287	0.064**	0.060*	0.074	0.068	34	0.271	0.082**	0.099**	0.110	0.100	31
		(0.029)	(0.035)	(0.056)	(0.062)			(0.040)	(0.047)	(0.077)	(0.090)	
		[3,019]	[2,585]	[1,372]	[1,176]			[1,454]	[1,245]	[668]	[530]	
5: BA attainment, in 6 years							0.635	0.061	0.058	0.101	0.134	30
								(0.041)	(0.049)	(0.077)	(0.090)	
								[1,454]	[1,245]	[668]	[512]	

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.
Notes: Robust standard errors are in the parentheses, sample size in brackets. The signs *, **, *** indicate significance at 10%, 5%, and 1% level, respectively. Explanatory variables not shown in the table are the running variable, an interaction between scholar status and the running variable, cohort dummies, interactions between cohort dummies and the running variable, interactions between cohort dummies, the running variable, and the scholar status, age, scaled GPA, ACT equivalent score, dummies for state of residence, a female dummy, ethnicity dummies, dummies for parental education, parents' income, free or reduced-lunch eligibility, receipt of food stamps, receipt of federal health insurance, receipt of Medicaid, an indicator for missingness of ACT score, and an indicator for missingness of food stamp receipt. To obtain the optimal bandwidth, we use a first-order polynomial, a uniform kernel, and bandwidth selector of Calonico et al. (2014a).

Table B.2: Sensitivity Analyses of RD estimates Using Different Optimal Bandwidths, 2009-2012

		2009-2012 Cohorts						
		Full	Intermediate	Narrow	Optimal	Optimal	Optimal	Optimal
Outcomes	μ	Sample	(±100)	(±40)	Bandwidth	Bandwidth	Bandwidth	Bandwidth
					(CCT, uniform)	(CCT, triangular)	(CCT, epanechnikov)	(IK, uniform)
1: Intermediate enrollment	0.853	0.015	0.022	0.023	0.014	0.027	0.024	0.041
		(0.022)	(0.025)	(0.037)	(0.036)	(0.032)	(0.033)	(0.039)
		[3,019]	[2,585]	[1,372]	[1,482]	[1,765]	[1,667]	[1,240]
2: 2^{nd} year persistence rate	0.752	0.023	0.036	0.032	0.025	0.033	0.042	0.052
		(0.026)	(0.030)	(0.044)	(0.043)	(0.038)	(0.040)	(0.046)
		[3,019]	[2,585]	[1,372]	[1,482]	[1,736]	[1,512]	[1,306]
3: 3^{rd} year persistence rate	0.674	0.046*	0.060*	0.081*	0.110**	0.082*	0.081*	0.113**
		(0.028)	(0.032)	(0.048)	(0.050)	(0.045)	(0.047)	(0.050)
		[3,019]	[2,585]	[1,372]	[1,240]	[1,512]	[1,425]	[1,278]
4: BA attainment, in 4 years	0.287	0.062**	0.056*	0.065	0.059	0.038	0.068	0.052
		(0.028)	(0.032)	(0.051)	(0.055)	(0.049)	(0.051)	(0.052)
		[3,019]	[2,585]	[1,372]	[1,176]	[1,482]	[1,392]	[1,342]

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.
Notes: Robust standard errors are in the parentheses and sample sizes are in brackets. The signs *, **, *** indicate significance at 10%, 5%, and 1% level, respectively. Explanatory variables not shown in the table are the running variable, an interaction between scholar status and the running variable, cohort dummies, interactions between cohort dummies and the running variable, interactions between cohort dummies, the running variable, and the scholar status, age, scaled GPA, ACT equivalent score, dummies for state of residence, a female dummy, ethnicity dummies, dummies for parental education, parents' income, free or reduced-lunch eligibility, receipt of food stamps, receipt of federal health insurance, receipt of Medicaid, an indicator for missingness of ACT score, and an indicator for missingness of food stamp receipt. To obtain the optimal bandwidth, we use a first-order polynomial, a uniform kernel, and bandwidth selector of Calonico et al. (2014a). The optimal bandwidth selector IK refers to that of Imbens and Kalyanaraman (2012).

Table B.3: Sensitivity Analyses of RD estimates Using Different Optimal Bandwidths, 2009-2010

Outcomes	μ	Full Sample	Intermediate ($\pm$100)	Narrow ($\pm$40)	Optimal Bandwidth (CCT, uniform)	Optimal Bandwidth (CCT, triangular)	Optimal Bandwidth (CCT, epanechnikov)	Optimal Bandwidth (IK, uniform)
1: Intermediate enrollment	0.850	0.032	0.037	0.064	0.098*	0.059	0.051	0.098*
		(0.031)	(0.036)	(0.054)	(0.057)	(0.052)	(0.052)	(0.057)
		[1,454]	[1,245]	[668]	[602]	[736]	[709]	[602]
2: 2^{nd} year persistence rate	0.751	0.034	0.054	0.072	0.103	0.067	0.075	0.067
		(0.037)	(0.043)	(0.064)	(0.067)	(0.059)	(0.060)	(0.061)
		[1,454]	[1,245]	[668]	[631]	[801]	[765]	[736]
3: 3^{rd} year persistence rate	0.658	0.067*	0.090*	0.121*	0.184**	0.132**	0.132*	0.244***
		(0.040)	(0.046)	(0.069)	(0.074)	(0.065)	(0.067)	(0.083)
		[1,454]	[1,245]	[668]	[602]	[736]	[709]	[502]
4: BA attainment, in 4 years	0.266	0.081**	0.096**	0.101	0.090	0.094	0.101	0.098
		(0.041)	(0.047)	(0.075)	(0.087)	(0.074)	(0.075)	(0.079)
		[1,454]	[1,245]	[668]	[530]	[696]	[668]	[619]
5: BA attainment, in 6 years	0.633	0.042	0.056	0.093	0.158*	0.122*	0.093	0.087
		(0.042)	(0.048)	(0.075)	(0.090)	(0.073)	(0.075)	(0.079)
		[1,454]	[1,245]	[668]	[512]	[723]	[677]	[619]

Source: the Dell Scholars Program Database, Michael and Susan Dell Foundation.
Notes: Robust standard errors are in the parentheses and sample sizes are in brackets. The signs *, **, *** indicate significance at 10%, 5%, and 1% level, respectively. Explanatory variables not shown in the table are the running variable, an interaction between scholar status and the running variable, cohort dummies, interactions between cohort dummies and the running variable, interactions between cohort dummies, the running variable, and the scholar status, age, scaled GPA, ACT equivalent score, dummies for state of residence, a female dummy, ethnicity dummies, dummies for parental education, parents' income, free or reduced-lunch eligibility, receipt of food stamps, receipt of federal health insurance, receipt of Medicaid, an indicator for missingness of ACT score, and an indicator for missingness of food stamp receipt. To obtain the optimal bandwidth, we use a first-order polynomial, a uniform kernel, and bandwidth selector of Calonico et al. (2014a). The optimal bandwidth selector IK refers to that of Imbens and Kalyanaraman (2012).

APPENDIX C

ADDITIONAL MATERIALS FOR CHAPTER 3

ADDITIONAL TABLES

Table C.1: The Effects of the Message on Expected Gains by Round: Range 30-60 Percent

	1: Round 1	2: Round 2	3: Round 3	4: Round 4
Message	-0.162*	-0.114	-0.007	-0.151
	(0.093)	(0.096)	(0.095)	(0.105)
Religiosity	0.028	0.038	-0.021	0.009
	(0.066)	(0.070)	(0.064)	(0.072)
Message · Religiosity	-0.097	-0.138	-0.214**	-0.229**
	(0.092)	(0.098)	(0.087)	(0.098)
Observations	243	243	243	243
Controls	Y	Y	Y	Y
Country FE	Y	Y	Y	N

Source: author's calculation using the experimental data.
Notes: standard errors are clustered at individual level. The signs *, **, *** indicates significance at 10, 5, and 1 percent. Other regression covariates not shown in the table are dummies for rounds, a dummy for gender, a dummy for China, standardized religiosity score, years in college, and standardized risk attitude score.

Table C.2: The Effects of Religious Message on Expected Gains by Consistent Subjects

	1: Criteria 1	2: Criteria 2	3: Criteria 3
Message	-0.096	-0.109	-0.107
	(0.086)	(0.083)	(0.076)
Religiosity	0.028	-0.001	-0.006
	(0.060)	(0.059)	(0.049)
Message · Religiosity	-0.133	-0.176**	-0.178**
	(0.084)	(0.083)	(0.073)
Observations	632	704	788
Controls	Y	Y	Y
Country FE	Y	Y	Y

Source: author's calculation using the experimental data.
Notes: standard errors are clustered at individual level. The signs *, **, *** indicates significance at 10, 5, and 1 percent. Other regression covariates not shown in the table are dummies for rounds, a dummy for gender, a dummy for China, standardized religiosity score, years in college, and standardized risk attitude score.

Table C.3: Summary Statistics of Muslim Lenders by Country

Variables	Indonesia		China	
	Baseline	Message	Baseline	Message
Gender, 1 if Male	0.48	0.32*	0.28	0.35
Age	21.30	21.30	19.40	20.50***
College year completed	3.22	3.22	2.22	2.75
Work Full or Part Time	0.23	0.27	0.28	0.38
Risk Attitude scale (Dospert)	56.21	50.76**	56.15	58.35
Religiosity scale	0.76	0.77	0.54	0.56
Observation	60	63	60	60

Source: author's calculation using the experimental data.
Notes: the signs *, **, *** indicates significance at 10, 5, and 1 percent.

ADDITIONAL FIGURES

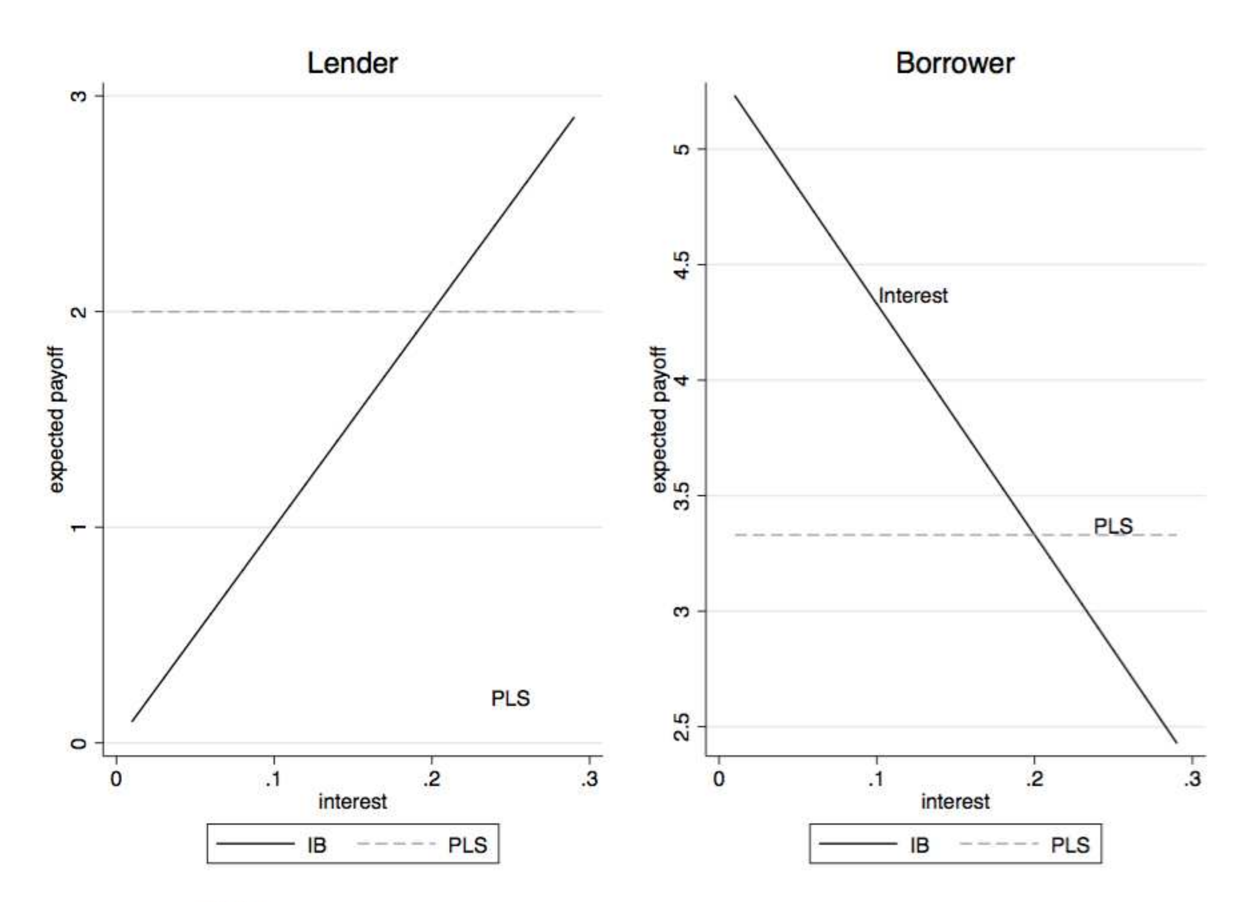

Figure C.1: Lender's and Borrower's Expected Payoffs for Interest Rate Range of 0-30 Percent

BIBLIOGRAPHY

Andreoni, J. and Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2):737–753.

Angrist, J., Hudson, S., Pallais, A., et al. (2014). Leveling up: Early results from a randomized evaluation of post-secondary aid. *NBER Working Paper No. 20800*.

Avery, C. (2013). Evaluation of the college possible program: Results from a randomized controlled trial. *NBER Working Paper No. 19562*.

Bailey, M. J. and Dynarski, S. M. (2011). Gains and gaps: Changing inequality in us college entry and completion. *NBER Working Paper No. 17633*.

Baum, S., Ma, J., and Payea, K. (2013). Education pays 2013. *The College Board*.

Bedi, A. S. and Edwards, J. H. (2002). The impact of school quality on earnings and educational returns—evidence from a low-income country. *Journal of Development Economics*, 68(1):157–185.

Benjamin, D. J., Choi, J. J., and Fisher, G. (2016). Religious identity and economic behavior. *Review of Economics and Statistics*, 98(4):617–637.

Bettinger, E. P. and Baker, R. B. (2014). The effects of student coaching: An evaluation of a randomized experiment in student advising. *Educational Evaluation and Policy Analysis*, 36(1):3–19.

Bettinger, E. P., Long, B. T., Oreopoulos, P., and Sanbonmatsu, L. (2012). The role of application assistance and information in college decisions: Results from the H&R block FAFSA experiment. *The Quarterly Journal of Economics*, 127(3):1205–1242.

Betts, J. R. (1995). Does school quality matter? evidence from the National Longitudinal Survey of Youth. *The Review of Economics and Statistics*, 77(2):231–250.

Betts, J. R. (1996). Do school resources matter only for older workers? *The Review of Economics and Statistics*, pages 638–652.

Betts, J. R. (2001). The impact of school resources on women's earnings and educational attainment: Findings from the national longitudinal survey of young women. *Journal of Labor Economics*, 19(3):635–657.

Blais, A.-R. and Weber, E. U. (2006). A domain-specific risk-taking (dospert) scale for adult populations. *Judgment and Decision Making*, 1(1).

Bratsberg, B. and Terrell, D. (2002). School quality and returns to education of US immigrants. *Economic Inquiry*, 40(2):177–198.

Brock, J. M., Lange, A., and Ozbay, E. Y. (2013). Dictating the risk: Experimental evidence on giving in risky environments. *The American Economic Review*, 103(1):415–437.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014a). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512):1753–1769.

Calonico, S., Cattaneo, M. D., Titiunik, R., et al. (2014b). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, 14(4):909–946.

Cameron, A. C. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372.

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics*, 3:1801–1863.

Card, D. and Krueger, A. (1994). The economic return to school quality: A partial survey. Technical report, Princeton University, Department of Economics, Industrial Relations Section.

Card, D. and Krueger, A. B. (1992a). Does school quality matter? returns to education and the characteristics of public schools in the United States. *The Journal of Political Economy*, 100(1):1–40.

Card, D. and Krueger, A. B. (1992b). School quality and black-white relative earnings: A direct assessment. *The Quarterly Journal of Economics*, 107(1):151–200 0033–5533.

Card, D. and Krueger, A. B. (1996). Labor market effects of school quality: Theory and evidence. *NBER Working Paper No. 5450.*

Carrell, S. E. and Sacerdote, B. (2013). Why do college going interventions work? *NBER Working Paper No. 19031.*

Castleman, B. and Goodman, J. (2016). Intensive college counseling and the enrollment and persistence of low income students. *Education Finance and Policy.*

Castleman, B. L. and Long, B. T. (2016). Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. *Journal of Labor Economics*, 34(4):1023–1073.

Castleman, B. L. and Page, L. C. (2014a). *Summer Melt: Supporting Low-Income Students through the Transition to College.* Harvard Education Press, Cambridge, MA.

Castleman, B. L. and Page, L. C. (2014b). A trickle or a torrent? understanding the extent of summer "melt" among college-intending high school graduates. *Social Science Quarterly*, 95(1):202–220.

Castleman, B. L. and Page, L. C. (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior & Organization*, 115:144–160.

Castleman, B. L. and Page, L. C. (2016). Freshman year financial aid nudges. *Journal of Human Resources*, 51(2).

Castleman, B. L., Page, L. C., and Schooley, K. (2014). The forgotten summer: Does the offer of college counseling after high school mitigate summer melt among college-intending, low-income high school graduates? *Journal of Policy Analysis and Management*, 33(2):320–344.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? evidence from Project STAR. *The Quarterly Journal of Economics*, 126(4):1593–1660.

Clotfelter, C. T., Hemelt, S. W., and Ladd, H. F. (2016). Multifaceted aid for low-income students and college outcomes: Evidence from North Carolina. *NBER Working Paper No. 22217*.

Condra, L. N., Isaqzadeh, M., and Linardi, S. (2017). Clerics and scriptures: Experimentally disentangling the influence of religious authority in Afghanistan. *British Journal of Political Science*, pages 1–19.

Dar, H. A. and Presley, J. R. (2000). Lack of profit loss sharing in Islamic banking: management and control imbalances. *International Journal of Islamic Financial Services*, 2(2):3–18.

Dearden, L., Ferri, J., and Meghir, C. (2002). The effect of school quality on educational attainment and wages. *Review of Economics and Statistics*, 84(1):1–20.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from head start. *American Economic Journal: Applied Economics*, 1(3):111–134.

DesJardins, S. L. and McCall, B. P. (2014). The impact of the Gates Millennium Scholars Program on college and post-college related choices of high ability, low-income minority students. *Economics of Education Review*, 38:124–138.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Dynarski, S. (2008). Building the stock of college-educated labor. *Journal of human resources*, 43(3):576–610.

Dynarski, S. M. (2003). Does aid matter? measuring the effect of student aid on college attendance and completion. *The American Economic Review*, 93(1):279–288.

Dynarski, S. M., Hemelt, S. W., and Hyman, J. M. (2015). The missing manual: Using national student clearinghouse data to track postsecondary outcomes. *Educational Evaluation and Policy Analysis*, 37(1_suppl):53S–79S.

El-Komi, M. and Croson, R. (2013). Experiments in islamic microfinance. *Journal of Economic Behavior & Organization*, 95:252–269.

El Massah, S. and Al-Sayed, O. (2013). Risk aversion and Islamic finance: an experimental approach. *European Journal of Applied Social Sciences Research*, 1(2).

Exley, C. L. (2016). Excusing selfishness in charitable giving: The role of risk. *The Review of Economic Studies*, 83(2):587–628.

Farooq, M. O. (2007). Partnership, equity-financing and Islamic finance: Whither profit-loss sharing? *Review of Islamic Economics (Special Issue)*, 11:67–88.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Goodman, J., Hurwitz, M., and Smith, J. (2015). College access, initial college choice and degree completion. *NBER Working Paper No. 20996*.

Grogger, J. (1996). Does school quality explain the recent black/white wage trend? *Journal of Labor Economics*, 14(2):231–253.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24(3):1141–1177.

Heckman, J., Layne-Farrar, A., and Todd, P. (1996). Human capital pricing equations with an application to estimating the effect of schooling quality on earnings. *The Review of Economics and Statistics*, 78(4):562–610.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the Econometric Society*, 47(1):153–161.

Howell, J. S. and Pender, M. (2016). The costs and benefits of enrolling in an academically matched college. *Economics of Education Review*, 51:152–168.

Hoxby, C., Turner, S., et al. (2013). Expanding college opportunities for high-achieving, low income students. *Stanford Institute for Economic Policy Research Discussion Paper*, (12-014).

Huber, S. and Huber, O. W. (2012). The centrality of religiosity scale (CRS). *Religions*, 3(3):710–724.

Hurwitz, M., Mbekeani, P. P., Nipson, M. M., and Page, L. C. (2016). Surprising ripple effects how changing the SAT score-sending policy for low-income students impacts college access and success. *Educational Evaluation and Policy Analysis*, 39(1):77–103.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 79(3):933–959.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.

Jacob, B. A. and Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *Review of Economics and Statistics*, 86(1):226–244.

Jana-Masri, A. and Priester, P. E. (2007). The development and validation of a Qur'an-based instrument to assess Islamic religiosity: The religiosity of Islam scale. *Journal of Muslim Mental Health*, 2(2):177–188.

Kane, T. J. (2003). A quasi-experimental estimate of the impact of financial aid on college-going. *NBER Working Paper No. 9703*.

Kneiding, C. and Rosenberg, R. (2008). Variations in microcredit interest rates. Technical Report 9510, The World Bank.

Krawczyk, M. and Le Lec, F. (2010). 'Give me a chance!' An experiment in social decision under risk. *Experimental Economics*, 13(4):500–511.

Lambarraa, F. and Riener, G. (2015). On the norms of charitable giving in islam: Two field experiments in morocco. *Journal of Economic Behavior & Organization*, 118:69–84.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355.

Ludwig, J. and Miller, D. L. (2007). Does head start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics*, 122(1):159–208.

Lynch, L. M. (1992). Private-sector training and the earnings of young workers. *The American Economic Review*, 82(1):299–312.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.

Miller, A. S. and Hoffmann, J. P. (1995). Risk and religion: An explanation of gender differences in religiosity. *Journal for the Scientific Study of Religion*, pages 63–75.

Nechyba, T. J. (1990). The southern wage gap, human capital and the quality of education. *Southern Economic Journal*, pages 308–322.

Noussair, C. N., Trautmann, S. T., Van de Kuilen, G., and Vellekoop, N. (2013). Risk aversion and religion. *Journal of Risk and Uncertainty*, 47(2):165–183.

Page, L. C. and Scott-Clayton, J. (2016). Improving college access in the United States: Barriers and policy responses. *Economics of Education Review*, 51:4–22.

Pallais, A. (2015). Small differences that matter: Mistakes in applying to college. *Journal of Labor Economics*, 33(2):493–520.

Perna, L. W. and Jones, A. (2013). *The state of college access and completion: Improving college success for students from underrepresented groups*. Routledge.

Rizzuto, R. and Wachtel, P. (1980). Further evidence on the returns to school quality. *Journal of Human Resources*, 15(2):240–254.

Rosenberg, R., Gaul, S., Ford, W., and Tomilova, O. (2013). Microcredit interest rates and their determinants: 2004–2011. In *Microfinance 3.0*, pages 69–104. Springer.

Sachs, N. (2009). Experimenting with identity: Islam, nationalism and ethnicity. In *APSA 2009 Toronto Meeting Paper*.

Schochet, P., Cook, T., Deke, J., Imbens, G., Lockwood, J., Porter, J., and Smith, J. (2010). Standards for regression discontinuity designs. *What Works Clearinghouse*.

Scott-Clayton, J. (2011). On money and motivation a quasi-experimental analysis of financial incentives for college achievement. *Journal of Human Resources*, 46(3):614–646.

Scrivener, S. and Weiss, M. J. (2013). More graduates: Two-year results from an evaluation of accelerated study in associate programs (ASAP) for developmental education students. *MDRC Policy Brief*.

Sellin, N. (1990). On aggregation bias. *International Journal of Educational Research*, 14(3):257–268.

Shariff, A. F. and Norenzayan, A. (2007). God is watching you: priming God concepts increases prosocial behavior in an anonymous economic game. *Psychological Science*, 18(9):803–809.

Shariff, A. F., Willard, A. K., Andersen, T., and Norenzayan, A. (2016). Religious priming: A meta-analysis with a focus on prosociality. *Personality and Social Psychology Review*, 20(1):27–48.

Strayer, W. (2002). The Returns to School Quality: College Choice and Earnings. *Journal of Labor Economics*, 20(3):475–503.

Tobias, J. L. and Li, M. (2003). A finite-sample hierarchical analysis of wage variation across public high schools: evidence from the NLSY and High School and Beyond. *Journal of Applied Econometrics*, 18(3):315–336.

Topel, R. H. and Ward, M. P. (1988). Job mobility and the careers of young men. *The Quarterly Journal of Economics*, 107(2):439–479.

Vella, F. (1998). Estimating models with sample selection bias: a survey. *Journal of Human Resources*, 33(1):127–169.

Wachtel, P. (1976). The effect on earnings of school and college investment expenditures. *The Review of Economics and Statistics*, 58(3):326–331.