Essays in Applied Microeconomics

Lesley Jeanne Turner

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012

ABSTRACT

Essays in Applied Microeconomics

Lesley Jeanne Turner

This dissertation broadly focuses on the role government should play in providing and financing education. The first chapter estimates the economic incidence of need-based student aid, one tool intended to ameliorate credit constraints in the market for higher education. The second chapter examines whether government programs providing support to low-income families should explicitly support or deny access to higher education by analyzing the impact of college attendance on the labor market outcomes of current and former welfare recipients. Chapter 3 focuses on publicly provided education at the primary and secondary levels, and estimates the impact of a teacher incentive pay program on student achievement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This dissertation would not have been possible without the support I have received from my committee, fellow students, and the faculty and staff of the Columbia's Economics Department. First, and foremost, I am deeply grateful to Miguel Urquiola, Jonah Rockoff, Bentley MacLeod, and Wojciech Kopczuk for their guidance and encouragement. Wojciech was incredibly generous with his time during the very crucial stages of preparing my job market paper. Bentley always pushed me to always think of the bigger picture. Miguel provided me with constant encouragement throughout graduate school. Jonah has been a wonderful mentor from the very beginning of my career as a researcher at Columbia and is a role model for the advisor I hope to be one day.

I consider myself especially lucky to have gone through Columbia's program with an amazing group of classmates. Reed Walker, Petra Persson, Ben Marx, and Todd Kumler met with me on a regular basis, providing me with much needed feedback on my ideas and research in its formative stages. Sarena Goodman is the coauthor of the third chapter of my dissertation and was a constant source of support during many unanticipated complications. Finally, I am very grateful to Afshin Zilanawala, Nicole Ngo, and Maya Rossin, for their encouragement and advice during many long hours at the School of Social Work.

My research benefited from insightful comments and discussions with many individuals and seminar participants. Thanks to Judy Scott-Clayton and Tom Bailey, I was able to present all three chapters of my dissertation in the Teacher's College Economics of Education seminar between 2009 and 2011. Numerous meetings with Till von Wachter helped make the first chapter of this dissertation a better paper. Additionally, I received valuable feedback on my first paper from Beth Akers, Stephanie Cellini, Janet Currie, Donald Davis, Yinghua He, Michael

Mueller-Smith, Christine Pal, Jim Sallee, Judy Scott-Clayton, Eric Verhoogen, and seminar participants at the U.S. Department of Treasury Office of Tax Analysis, MDRC, the Upjohn Institute, the University of Notre Dame, the University of Illinois at Urbana-Champaign, the Stanford Institute for Economic Policy Research, the University of Pittsburgh, the Federal Reserve Board of Governors, the University of Maryland at College Park, the University of Toronto, the Wharton School of the University of Pennsylvania, the University of Wisconsin at Madison, Brown University, and the 2012 Association for Education Finance and Policy (AEFP) conference. My second paper benefited from suggestions provided by participants at the 2010 Association for Public Policy Analysis and Management conference, the 2011 Colorado Works Professional Development Academy, and the 2011 Welfare Research and Evaluation Conference. Finally, on my third paper, Sarena and I are grateful for the feedback we received from Derek Neal, Jesse Rothstein, Till von Wachter, and participants at the 2010 AEFP conference and the 2010 Harvard Kennedy School's Program on Education Policy and Governance's conference.

This dissertation truly was made possible through the help of numerous individuals and organizations that provided me with data. Tom Bailey and the Columbia Community College Research Center generously providing me with access to the NPSAS data used in the first chapter of this dissertation, while Matt Zeidenburg provided invaluable assistance in understanding and using these files. I am grateful to Sarah Marsh, Vince Romeo, Jessica Valand and the Colorado Department of Human Services, Colorado Department of Higher Education, and Colorado Department of Labor and Employment for the data used in my dissertation's second chapter. Finally, the New York City Department of Education generously provided data on schools participating in the city's bonus program experiment.

My research was supported by grants from the American Education Research Association, which receives funds for its "AERA Grants Program" from the National Science Foundation, and Eunice Kennedy Shriver National Institute of Child Health and Human Development. Finally, I am grateful to the Columbia Population Research Center for providing a wonderful environment with excellent colleagues and to Stephanie Cellini and the George Washington Institute of Public Policy for housing me during a crucial summer.

Finally, my husband Josh Rosenfeld provided endless and unflinching support throughout my time at Columbia. His encouragement, humor, and love made all the difference.

To my parents, for their unconditional love,
and to my grandmother Laura,
whose courage helped me find my own.

INTRODUCTION

In the United States, like many developed countries, government provides substantial support for primary, secondary, and postsecondary education. In 2011, expenditures on education equaled 7.6 percent of GDP (National Center for Education Statistics 2011). At the postsecondary level, federal grants and loans for college students aim to ameliorate credit market imperfections. Publicly provided K-12 education insures universal access to human capital development for all children. This dissertation broadly addresses the question what role government should play in financing and providing education.

The first chapter estimates the economic incidence of need-based student aid, one tool used to address credit market imperfections in the market for higher education. The federal Pell Grant Program provides billions of dollars in subsidies to low-income college students to increase affordability and access to higher education. Using regression discontinuity (RD) and regression kink (RK) designs, I show that 16 percent of all Pell Grant aid is passed-through to schools in the form of higher effective prices. Additionally, I reconcile differences between RD and RK estimates using a framework in which the treatment of Pell Grant aid is multidimensional: students receive an additional dollar of Pell Grant aid and are also labeled as Pell Grant recipients. RD estimates confound the effects of these dimensions, which have opposite impacts on schools' pricing decisions. With my combined RD/RK approach, I am able to separately identify schools' willingness to pay for students categorized as needy and the pricing response to outside subsidies.

The second chapter of my dissertation examines whether government programs providing support to low-income families should explicitly support or deny access to higher education. I estimate the impact of college credits and credentials on the labor market outcomes of several cohorts of current and former welfare recipients. I use an event-study approach to control for time-invariant individual characteristics, such as differences in ability and motivation. Women who are induced to attend college after entering welfare experience large and significant earnings gains, however, these returns are driven by credential receipt and when sub-associate's degree credentials are unobservable, positive earnings gains will be inappropriately attributed to college attendance alone.

The final chapter of my dissertation addresses the question of how to efficiently provide education at the primary and secondary levels. Teacher compensation schemes are often criticized for lacking a performance-based component. Proponents argue that teacher incentive pay can raise student achievement and stimulate system-wide innovation. My coauthor and I examine a group-based teacher incentive scheme implemented in New York City and investigate whether specific features of the program contributed to its ineffectiveness. Although overall the program had little effect on student achievement, we show that in schools where incentives to free-ride were weakest, the program led to small increases in math achievement. Our results underscore the importance of carefully considering the design of teacher incentive pay programs.

CHAPTER 1

The Incidence of Student Financial Aid: Evidence from the Pell Grant Program

1.1 Introduction

The federal government provides billions of dollars in targeted need-based aid to low-income college students every year. Although students are the statutory recipients of this aid, its economic incidence may fall partially on schools (Fullerton and Metcalf 2002). Specifically, schools may strategically increase recipients' effective prices, crowding out federal aid by reducing discounts provided through institutional aid. Concurrent tuition and student aid increases combined with substantial growth in the for-profit sector of higher education underscore the importance of evaluating federal aid crowd out.

In this paper, I estimate the economic incidence of the federal Pell Grant Program, the largest source of need-based student aid in the United States, using detailed student-level data from the National Postsecondary Student Aid Study. I show that institutions capture 16 percent of all Pell Grant aid – approximately $6 billion in 2011 – through price discrimination. Furthermore, I illustrate that the extent and pattern of capture vary substantially by institutional control and selectivity. For example, on average, public schools do not capture any Pell Grant aid, while among students attending selective nonprofit schools, decreases in institutional aid crowd out over 50 percent of the value of Pell Grant aid. Additionally, the incidence of the Pell Grant Program also varies across students within some sectors. For instance, among public school students near the Pell Grant eligibility threshold, Pell Grant aid appears to *crowd in* rather than crowd out institutional aid.

I identify these impacts using discontinuities in the relationship between Pell Grant aid and the federal government's measure of student need. Specifically, the Pell Grant Program's schedule contains discontinuities in both the level and in the slope of aid, resulting in students with very similar levels of need receiving significantly different grants. This variation allows me to use both regression discontinuity (RD) and regression kink (RK) designs (Angrist and Lavy 1999; Hahn et al. 2001; Card et al. 2009; Nielsen et al. 2010). My analysis illustrates the relationship between these two methods and provides an example of circumstances under which RD and RK designs will yield significantly different conclusions.

The RK approach relates the change in the slope of the Pell Grant schedule at the eligibility cut-off with the change in the slope of the institutional aid schedule at this same point. RK estimates imply that, on average, schools capture a portion of Pell Grant aid through price discrimination. In contrast, the RD approach relates the change in the level of Pell Grant aid at the eligibility cut-off with the change in the level of institutional aid at this same point. RD estimates imply that, on average, schools *increase* institutional aid for Pell Grant recipients.

I reconcile these disparate estimates using a framework in which the “treatment” of Pell Grant receipt is multidimensional. Specifically, students at the margin of Pell Grant eligibility receive an extra dollar of outside aid but are also given the label of being a Pell Grant recipient, which may change some institutions’ willingness to direct resources towards them. I show that it is possible to identify both schools’ willingness to pay for students categorized as Pell Grant recipients and their pricing response to outside subsidies using a combined RD/RK approach.

RD estimates only identify the combined impact of these dimensions, and near the Pell Grant eligibility threshold, a greater willingness to pay dominates the pass-through of outside aid from students to schools. This result is misleading, however, since using my combined RD/RK approach, I estimate that fewer than one third of Pell Grant recipients benefit from these transfers. This is because the pass-through of each additional dollar of Pell Grant quickly overtakes these schools’ willingness to pay for needy students. My results suggest that, on average, Pell Grant recipients receive an additional $260 in institutional aid due to schools’ willingness to pay for needy students, but every additional dollar of Pell Grant aid is crowded out by a 22 cent reduction in institutional aid.

My paper is one of the first to combine these two identification strategies and the first to explicitly show how a multidimensional treatment affects RD estimates. Although the Pell Grant Program provides an especially stark example of how a multidimensional treatment affects RD estimates, in other circumstances where both a discontinuity and a kink are present, my results suggest that additional information can potentially be gained from using my combined RD/RK approach.

Little is known about how institutions compete for students or the objectives of public and nonprofit schools. This paper provides insight into the industrial organization of higher education by showing how variation in schools' response to Pell Grant aid relates to differences in schools' objectives and market power across sectors. Public schools demonstrate a willingness to pay for students categorized as Pell Grant recipients. Although the net capture of Pell Grants in the public sector is close to zero, increases in institutional aid for recipients near the eligibility threshold come at the expense of the neediest Pell recipients. Conversely, selective nonprofit institutions appropriate 79 percent of their students' Pell Grants, suggesting these schools have considerable market power.

The for-profit sector of higher education has grown substantially over the last decade and in recent years, has been criticized for unethical marketing practices and financial aid fraud (U.S. Government Accountability Office 2010). Although these schools disproportionately serve federal aid recipients, I find that for-profit institutions behave no differently than nonselective nonprofit schools and, combined, these schools appropriate only 18 percent of their students' Pell Grants.

Finally, this paper contributes to a broader literature on the effectiveness of targeted subsidies and the importance of considering impacts on the behavior of both consumers and firms.[1] Research by Long (2004) and Turner (2012) suggests that other sources of financial aid crowd out institutional discounts by as much as 100 percent. However, previous studies specifically focusing on the Pell Grant Program estimate a positive correlation between prices and Pell Grant generosity. These impacts are identified using time-series variation in the maximum award, variation that is likely correlated with unobservable year specific shocks to the economy (e.g., McPherson and Schapiro 1991; Singell and Stone 2007).

The remainder of this paper proceeds as follows: the next section describes the Pell Grant program and previous estimates of the impact of student aid on prices. Section 1.3 discusses the NPSAS data and presents descriptive statistics, while Section 1.4 describes the regression kink design and my estimation strategy. Section 1.5 presents results from RD and RK estimates and Section 1.6 provides a conceptual framework that reconciles differences between these estimates. I estimate the overall incidence of the Pell Grant Program in Section 1.7, while Section 1.8 concludes.

1.2 The Pell Grant Program and Need-based Student Aid

An extensive literature estimates large private returns to higher education and positive externalities associated with a highly educated population.[2] Between 1979 and 2009, real tuition and fees increased by close to 200 percent, outpacing growth in income and student aid (National Center for Education Statistics, 2011). If some individuals face credit constraints and cannot

---

[1] For example, Rothstein (2008) shows that Earned Income Tax Credit (EITC) induced increases in labor supply drive down wages, and firms receive over half of the benefit of EITC payments. Hastings and Washington (2010) show that grocery stores benefit from public assistance via cyclical pricing in response to recipients' impatience.

[2] For example, see Card (1999), Moretti (2004), Lochner and Moretti (2004), and Dee (2004).

borrow against future income to finance college attendance, education levels may be inefficiently low. For these reasons, the United States federal and state governments provide substantial subsidies to low-income college students.

Established to promote access to postsecondary education, the Pell Grant Program is the largest source of need-based student aid in the United States. In 2011, the program provided 9.5 million low-income students with subsidies totaling $35 billion. The maximum Pell Grant award has grown in generosity from $452 during the 1973-1974 school year (hereafter, 1974) to $5,550 in 2011, a 62 percent increase in real terms (Figure 1.1). However, over this period the purchasing power of the maximum Pell Grant award declined (Figure 1.2). In 2010, the maximum Pell Grant represented 42 percent of average tuition and fees at public institutions and only 17 percent at private schools (National Center for Education Statistics, 2011).

A student's Pell Grant award depends both on the annual maximum award and upon her expected family contribution (EFC), the federal government's measure of need. Students are required to complete a Free Application for Federal Student Aid (FAFSA) to qualify for Pell Grants and other federal student aid (e.g., loans, work-study). FAFSA inputs include a detailed set of financial and demographic information, such as income, untaxed benefits, assets, family size and structure, and number of siblings in college. When filing the FAFSA, students also must specify up to (but no more than) six schools they are considering attending.[3]

The federal government calculates a student's EFC using a complicated, non-linear function of these inputs (e.g., U.S. Department of Education 2006). The federal government provides the listed schools with the student's EFC and FAFSA inputs, and these schools calculate federal (and in some cases state) grants and loans. With this information in hand,

[3] Beginning in 2009, students could specify up to 10 schools that would receive their FAFSA information.

schools choose how institutional aid will be distributed across students. Thus, a school observes the student's FAFSA, EFC, and outside aid before deciding the level of its own discount, which it provides via institutional aid. Students receive a financial aid package from each school specifying federal, state, and institutional grant aid and loans. Students do not observe their Pell Grant award until this point, where it is included as a component of the final price displayed in their financial aid package.

A full-time, full-year student i in year t qualifies for a Pell award equal to:

$$(1.1) \qquad Pell_{it} = \left(maxPell_t - EFC_{it}\right)\cdot \mathbf{1}\left[maxPell_t - EFC_{it} \geq 400\right] + 400\cdot \mathbf{1}\left[maxPell_t - EFC_{it} \in \left(400, 200\right]\right]$$

Where $maxPell_t$ is the maximum Pell award available in year t (Figure 1.1), EFC_{it} is the expected family contribution of student i in year t, and $\mathbf{1}[\cdot]$ is the logical indicator function. Pell Grant awards are rounded up to the next \$100 and students qualifying for an award between \$399 and \$200 receive \$400.[4] Students who qualify for less than \$200 in aid do not receive a Pell Grant. The Pell Grant formula generates two sources of variation that I use for identification. First, crossing the Pell Grant eligibility threshold leads to a discrete increase in a student's statutory award, from \$0 to \$400, which enables me to use a regression discontinuity design. Second, the variation created by the change in the slope of the Pell Grant-EFC function, from 0 to -1, allows me to use a regression kink design.[5] Figure 1.3 displays the Pell Grant award schedule in 1996, 2000, 2004, and 2008.[6]

---

[4] The minimum Pell Grant award increased to \$890 in 2009, \$976 in 2010, and \$1176 in 2011. However, the minimum award remained \$400 in the years I examine.

[5] The Pell Grant formula for part-time students is $Pell_{it} = \max\{(0.5 * maxPell_t - EFC_{it}), 0\}$; the change in the slope of the relationship between EFC and Pell Grant aid is -0.5. The minimum Pell Grant does not depend on attendance intensity. Part-year students receive a prorated Pell Grant.

[6] Eligibility for other types of federal aid (e.g., Supplemental Educational Opportunity Grants, Stafford loans, work study) also depends on the EFC. The Pell Grant Program is the only federal entitlement for college students.

1.2.1 Previous Estimates of the Impact of Pell Grant Aid on College Enrollment and Prices

Tuition and financial aid influence important outcomes, from the decision to enroll in college, to persistence and degree completion (Angrist 1993; Bound and Turner 2002; Dynarski 2003; Bettinger 2004). Although the Pell Grant Program aims to increase low-income students' access to higher education, past research finds little impact on college enrollment except for older, non-traditional students (Kane 1995; Seftor and Turner 2002, Deming and Dynarski 2010). Students only receive information concerning the level of their Pell Grant after they have submitted a FAFSA, and this information is provided as part of a school's financial aid package, where the final price (tuition net of state, federal, and institutional grants) is likely the most salient feature. If low-income students lack information about the Pell Grant Program, Pell Grant aid may not increase college enrollment. The complexity of the FAFSA form imposes a large cost on potential students (Dynarski and Scott-Clayton 2008). Bettinger et al. (forthcoming) show that information on the availability of financial aid and assistance with the FAFSA application process increase the likelihood of enrollment.

The relatively weak response of student demand to Pell Grant aid suggests the potential for schools to appropriate student aid through price increases. However, previous studies show no conclusive evidence that increases in Pell Grant generosity cause schools to raise prices. McPherson and Schapiro (1991) show that overall institutional aid levels are positively correlated with Pell Grant generosity; Singell and Stone (2007) find a positive correlation between published tuition and Pell Grant generosity among private institutions. In both cases, identification comes from time-series variation in the maximum Pell Grant.

Raising tuition is only one method schools may use to benefit from Pell Grant generosity. Schools can also adjust students' prices through price discrimination by reducing institutional aid. The practice of price discrimination, or offering a schedule of prices that varies according to consumer demand elasticities, has been documented in a variety of imperfectly competitive markets and the market for higher education is unique in the extensive amount of customer information schools observe, including a measure of students' ability to pay.[7]

Two studies explicitly examine whether student aid crowds out institutional aid. Turner (2012) estimates the incidence of education tax credits, which primarily benefit middle-income students, and finds that schools reduce institutional aid dollar for dollar as tax-based aid increases. Long (2004) examines the implementation of the Georgia HOPE scholarship program, which provides substantial assistance to students in Georgia who achieve a 3.0 GPA. Public schools responded to the HOPE program by increasing fees, capturing 10 percent of HOPE aid, while private nonprofit institutions captured 30 percent of HOPE aid by increasing tuition and fees and reducing institutional aid. Additionally, using administrative Pell Grant data and a simulated instrumental variables approach, Li (1999) finds a positive relationship between Pell Grant aid and both listed tuition and net tuition per student. By comparing the impact of Pell Grant aid on net tuition per student and listed tuition per student, it is possible to infer whether schools also alter institutional aid. Results suggest that four-year institutions both increase tuition and reduce institutional aid in response to Pell Grant generosity

Schools' response to the HOPE program suggests that the economic incidence of the Pell Grant Program may vary by institutional control. Traditionally, public and nonprofit schools primarily serve the market for higher education. However, the last decade has seen substantial growth in the for-profit sector. For-profit institutions increasingly serve low-income students and have been criticized for high student loan default rates and deceptive recruiting practices (U.S.

[7] E.g., housing (Yinger 1998), loans (Charles et al. 2008), and vehicles (Langer 2011).

Government Accountability Office 2010) and charge significantly higher tuition than comparable public schools.[8] Influenced by these concerns, "gainful employment" legislation will specifically regulate programs primarily offered by for-profit schools beginning in 2012.[9]

1.3 Data and Descriptive Statistics

The National Postsecondary Student Aid Study (NPSAS) is a restricted-use, nationally representative, repeated cross-section of college students.[10] I observe each student's EFC, demographic characteristics, FAFSA inputs (e.g., family income and assets), and financial aid provided by the federal government and other sources. My sample includes students present in the 1996, 2000, 2004, and 2008 NPSAS waves. I eliminate graduate and first-professional students as well as noncitizens and non-permanent residents, as these students are ineligible for Pell Grant aid. Additionally, I exclude students who attended multiple schools in the survey year, received athletic scholarships, and were not enrolled in the fall semester[11]

I exclude all students attending schools that only offer sub-associate certificate programs, theological seminaries, and other faith-based institutions, since many of these schools are not eligible to distribute federal aid. Finally, I focus on students whose EFC is within $10,000 of the Pell Grant eligibility threshold, although in a subset of analyses, I look at students within

[8] The share of Pell Grant recipients attending for-profit schools increased from 13 to 25 percent between 2000 and 2010 (Pell Grant End of Year Reports). Conversely, the share of all students at for-profit schools grew from 4 to 11 percent (Deming et al. 2012). The 2009, 15 percent of former for-profit students defaulted on their student loans within two years of exiting college. The rates for public and non-profit institutions were 7 and 5 percent, respectively. In 2010, average public school tuition was $5,000; for-profit students paid $15,700 (National Center for Education Statistics 2011).

[9] The legislation requires that for-profit institutions and certificate programs in other sectors prepare students for "gainful employment" to qualify for federal student aid (76 FR 34386).

[10] After the original 2008 NPSAS sample was drawn, additional observations of National Science and Mathematics Access to Retain Talent (SMART) Grant recipients were added. For my main set of estimates, I drop oversampled SMART Grant recipients. My results are robust to using the NPSAS sample weights and retaining SMART Grant recipients or excluding observations from 2008, the first year of the NPSAS in which students eligible for SMART Grants could potentially be sampled (results available upon request).

narrower windows around this threshold. My final sample includes approximately 133,270 undergraduate students attending 1,800 unique institutions. Due to the National Center for Education Statistics' confidentiality requirements, all NPSAS sample sizes are rounded to the nearest 10 observations.

My sample includes new and continuing students. Although upper-year students likely have less elastic demand than first year students, EFC and institutional aid are highly correlated over time. Schools award multi-year institutional aid packages and for many students, one of the primary components of EFC – family income – does not vary substantially over time.

I classify schools by selectivity and control, distinguishing between public, nonprofit, and for-profit institutions that are either selective or nonselective. I use the IPEDS and Barrons' Guide to determine an institution's selectivity. The IPEDS contains annual data on acceptance rates and the Barrons' College Guide classifies four-year public and nonprofit institutions into six categories of selectivity based on acceptance rates, college entrance exam performance, and the minimum class rank and grade point average required for admission. First, I classify all for-profit and institutions offering two-year programs as non-selective. If the IPEDS lists an institution as "inclusive" (i.e., open admissions), I also classify it as nonselective. Finally, I classify remaining institutions as nonselective if either the Barrons' Guide lists them as less competitive or non-competitive or they are missing Barrons' rankings and admit more than 75% of applicants. Under this scheme, schools I classify as "selective" are not highly selective. Rather, these schools reject some portion of applicants.[12]

Public schools are either operated by publicly elected or appointed officials or receive the majority of their funding from public sources. Conversely, private institutions receive the

[12] On average, selective public institutions admit 61 percent of applicants and selective nonprofits admit 56 percent.

majority of funding from private sources and are run by privately appointed individuals. Nonprofit institutions are exempt from federal taxes but are subject to the "non-distribution constraint" which prohibits a school from distributing revenue to its controlling body in excess of regular wages and other operating expenses (Hansmann 1980).[13] For-profit schools pay corporate income taxes, but may also distribute profits to owners or shareholders.

1.3.1 Characteristics of Students and Schools

Table 1.1 displays summary statistics by sector. In my sample, for-profit students are the most likely to receive Pell Grant aid, while students attending selective institutions are the least likely to receive Pell Grants. However, conditional on receiving a Pell Grant, award amounts are similar across sectors and approximately three quarters of Pell Grant recipients receive less than the maximum award. Schools in all sectors use institutional aid to provide discounts from the list price, although students attending for-profit and nonselective public institutions are the least likely to receive these discounts. On average, for-profit students are more likely to be non-white and are older than students in other sectors. Students attending nonselective schools are more likely to be classified as independent, a status given to students who are married, have dependents, are veterans, or are older than 24.

I use information from the Integrated Postsecondary Student Data System (IPEDS) to examine overall revenue and expenditures for schools in my sample. The IPEDS contains the universe of institutions that receive federal student aid and the U.S. Department of Education collects detailed information on school characteristics, enrollment, faculty and staff, and finances through annual surveys. Table 1.2 displays institutional revenue and expenditures for each sector

[13] Internal Revenue Code (IRC) section 501(c)(3). Income from activities unrelated to the provision of education is subject to taxation (IRC sections 511-514).

using data from the IPEDS. For-profit schools are the only institutions that receive a substantial portion of revenue from the Pell Grant Program (14 percent). Pell Grants represent only 7 percent of public nonselective schools' revenue, 4 percent of revenue in private nonselective schools, and only 1 to 2 percent for more selective institutions. With the exception of for-profit institutions, these calculations suggest that relative to other sources of revenue, even if institutions responded to Pell Grant increases by raising overall tuition, the potential gains would be quite small.

1.4 Empirical Framework

Previous studies identify the impact of Pell Grant aid on prices using time series variation in the maximum award. However, if aid generosity is correlated with unobservable time-varying shocks, these estimates will suffer from omitted variables bias. Since Pell Grant generosity also varies across students within a given year, it is possible to separate the impact of Pell Grant aid from year-specific shocks under the assumption that, conditional on observables, Pell Grant aid is not correlated with unobservable student characteristics. This is a strong assumption. Pell Grant generosity is increasing in need, and while I can explicitly control for EFC, the specific functional form of the relationship between EFC and unobservable heterogeneity is unknown.

To overcome concerns of omitted variables bias, I take advantage of the relationship between Pell Grant aid and EFC. Specifically, I identify the impact of Pell Grant aid on student prices using variation induced by the kink and the discontinuity in the relationship between Pell Grant and EFC. The kink occurs where the slope of the $Pell(efc)$ schedule changes from 0 to -1, while the discontinuity is driven by the increase from in Pell Grant aid from \$0 to \$400 at the eligibility threshold, due to the rounding-up of awards scheduled to fall between \$200 and \$400. This variation allows me to use both a regression discontinuity (Angrist and Lavy 1999; Lee and

Lemieux 2010) and a regression kink design (Card et al. 2009; Nielsen et al. 2010). Like the regression discontinuity design, the regression kink estimator identifies the average treatment effect for individuals near the eligibility cut-off under specific conditions.

1.4.1 Regression Kink and Regression Discontinuity Designs

Similar to the regression discontinuity (RD) design, the regression kink (RK) design allows for identification of the impact of an endogenous regressor that is a known function of an observable assignment variable (Card et al. 2009). Here, the endogenous regressor is Pell Grant aid, while EFC is the assignment variable. The RK design uses variation induced by a change in the slope of the relationship between Pell Grant aid and EFC as the eligibility threshold is approached from above and below. Like the RD design, the RK design will be invalidated if individuals are able to sort perfectly in the neighborhood of the kink.

Let $Y = f(Pell, \tau) + g(EFC) + U$ represent the causal relationship between institutional aid, *Y*, and Pell Grant aid, $Pell = Pell(EFC)$, in a given school and year, where *U* is random vector of unobservable, predetermined characteristics. Given the existence of a kink in the Pell Grant schedule, the required identifying assumptions are: (1) the direct marginal impact of EFC on institutional aid is continuous and (2) the conditional density of EFC (with respect to *U*) is continuously differentiable at the threshold for Pell Grant eligibility (Card et al. 2009). These assumptions encompass those required for identification using a RD design, which requires institutional aid to be continuous (rather than continuously differentiable) in EFC and that the conditional density of EFC be continuous (rather than continuously differentiable). Essentially, even if many other factors affect college pricing decisions, as long as there are no discontinuities in the relationship between these factors and EFC at the eligibility threshold, the RK estimator approximates random assignment in the neighborhood of the kink. Additionally, as in the case of the RD design, the second assumption generates testable predictions concerning how the density of EFC and the distribution of observable characteristics should behave in the neighborhood of the eligibility cut-off.

Assume that each additional dollar of Pell Grant aid has the same marginal effect on schools' pricing decisions (at least in the neighborhood of the eligibility threshold):

$$f(Pell,\tau) = \tau_1 Pell \tag{1.2}$$

In this case, τ_1 represents the pass-through of each additional dollar of Pell Grant aid from students to schools.

If the required identifying assumptions hold, the RK estimator identifies:

$$\tau_{RK} = \frac{\lim\limits_{\varepsilon\uparrow 0} \frac{\partial E[Y \mid EFC = efc_0 + \varepsilon]}{\partial efc} - \lim\limits_{\varepsilon\downarrow 0} \frac{\partial E[Y \mid EFC = efc_0 + \varepsilon]}{\partial efc}}{\lim\limits_{\varepsilon\uparrow 0} \frac{\partial E[Pell \mid EFC = efc_0 + \varepsilon)]}{\partial efc} - \lim\limits_{\varepsilon\downarrow 0} \frac{\partial E[Pell \mid EFC = efc_0 + \varepsilon)]}{\partial efc}} = \tau_1 \tag{1.3}$$

Where efc_0 represents the eligibility threshold for the Pell Grant Program. Since the Pell Grant Program's schedule also contains a discontinuity in the *level* of aid at the eligibility threshold, I can also identify the impact of Pell Grant aid on college pricing decisions using a RD design. As long as equation (1.2) describes the relationship between Pell Grant aid and colleges' pricing decisions, the RD estimator also identifies τ_1:

$$\tau_{RD} = \frac{\lim\limits_{\varepsilon\uparrow 0} E[Y \mid EFC = efc_0 + \varepsilon] - \lim\limits_{\varepsilon\downarrow 0} \partial E[Y \mid EFC = efc_0 + \varepsilon]}{\lim\limits_{\varepsilon\uparrow 0} E[Pell \mid EFC = efc_0 + \varepsilon)] - \lim\limits_{\varepsilon\downarrow 0} E[Pell \mid EFC = efc_0 + \varepsilon)]} = \tau_1 \tag{1.4}$$

In practice, my estimation strategy involves "fuzzy" RD/RK. Some eligible students do not apply for federal aid and thus, do not receive Pell Grants.[14] Additionally, variables in the NPSAS contain measurement error induced by random perturbations to preserve respondent

[14] Results are robust to eliminating students who do not submit a FAFSA (available upon request).

confidentiality. Since the location of the Pell Grant Program's eligibility threshold changes as the maximum award increases, I create a standardized measure of the distance of a student's EFC from the year-specific EFC representing the eligibility threshold: $E\tilde{F}C_{it} = EFC_{it} - efc_{0t}$, where efc_{0t} is the cut-off for Pell Grant eligibility in year *t* and all students with $E\tilde{F}C_{it} \geq 0$ are ineligible for Pell Grant aid.[15] Figure 1.4 displays the empirical distribution of Pell Grant aid for students in my sample by standardized EFC.[16]

Consider the following first stage and reduced form equations:

$$(1.5) \qquad Pell_{it} = \eta \cdot \mathbf{1}\left[E\tilde{F}C_{it} < 0\right] + \lambda \cdot \mathbf{1}\left[E\tilde{F}C_{it} < 0\right] \cdot \left(E\tilde{F}C_{it}\right) + \sum_{\rho}[\psi_{\rho} \cdot \left(E\tilde{F}C_{it}\right)^{\rho}] + \theta_j + \theta_t + \upsilon_{ijt}$$

$$(1.6) \qquad y_{ijt} = \beta \cdot \mathbf{1}\left[E\tilde{F}C_{it} < 0\right] + \gamma \cdot \mathbf{1}\left[E\tilde{F}C_{it} < 0\right] \cdot \left(E\tilde{F}C_{it}\right) + \sum_{\rho}[\pi_{\rho} \cdot \left(E\tilde{F}C_{it}\right)^{\rho}] + \delta_j + \delta_t + \varepsilon_{ijt}$$

Here, *i* indexes students, *j* indexes institutions, and *t* indexes years. $Pell_{it}$ is the Pell Grant award received by student *i* in year *t*, while y_{ijt} represents the institutional aid provided by school *j* to this student. The term $\mathbf{1}\left[E\tilde{F}C_{it} < 0\right]$ is an indicator for Pell Grant eligibility while ρ indexes the degree of polynomial in the assignment variable, $E\tilde{F}C_{it}$. I include year and school fixed effects as well as a vector of student characteristic to reduce residual variation; these terms are not necessary for identification.[17] The ratio of the reduced form and first-stage coefficients for the

---

[15] For the years I examine, efc_{0t} equals \$2140 (1996), \$2925 (2000), \$3850 (2004), and \$4110 (2008).

[16] In a given year, the kink and discontinuity in the relationship between Pell Grant aid and EFC occur at slightly different values of EFC (see Appendix Figure A2). However, the distance between these points is quite small and only a small fraction of students have an EFC falling on this "plateau". I treat both the slope and the level of Pell Grant funding changes as occurring at the eligibility cut-off. My results are robust to removing students whose EFC falls on the plateau (forcing the discontinuity and kink to occur at the same value of EFC).

[17] These characteristics include indicators for gender, race, fall attendance intensity, enrollment length, level (e.g., whether the student is a first year, second year, etc.), out-of-state student, and a quadratic in age.

interaction between $\mathbf{1}\left[E\tilde{F}C_{it} < 0\right]$ and the linear term in $E\tilde{F}C_{it}$, $\hat{\tau}_{RK} = \frac{\hat{\gamma}}{\hat{\lambda}}$, represents the RK estimate of the impact of Pell Grant aid on institutional aid. Likewise, the ratio of the reduced form and first-stage coefficients for Pell Grant eligibility, $\hat{\tau}_{RD} = \frac{\hat{\beta}}{\hat{\eta}}$, represents the RD estimate of the impact of Pell Grant aid on institutional aid.

To further illustrate the mechanics of this framework, Figures 1.5A and 1.5B illustrate potential behavior of the relationship between institutional aid and EFC near the eligibility threshold (i.e., potential values for γ and β) and corresponding implications for RK and RD estimates. In the first case (Figure 1.5A), there is no discontinuity or kink in the relationship between institutional aid and EFC near the eligibility cut-off – the change in the level and the slope of institutional aid are both equal to zero – indicating students receive the full benefit of Pell Grant aid. In this case, the RD and RK estimators both yield an estimate of zero. Conversely, Figure 1.5B illustrates the case where Pell Grant aid fully crowds out institutional aid. The change in the level of institution aid is equal (in absolute value) to the change in the level in Pell Grant aid, suggesting the RD approach will yield an estimate of -1, or full pass-through of Pell Grant aid from students to institutions. The change in the slope of institutional aid at the eligibility threshold is likewise equal (in absolute value) to the change in the slope of the Pell Grant schedule at this same point, also resulting in an estimate of -1. As in the first example, both RD and RK designs produce the same estimate; in this case, suggesting schools appropriate 100 percent of Pell Grant aid.

1.4.2 Evaluating the RD and RK Identifying Assumptions

Identification using the RK or RD design hinges on the assumption that students cannot exactly sort to obtain a more advantageous EFC. Students and their parents likely act to reduce their estimated need, but as long as they cannot chose an exact value of EFC, the RK and RD estimators will be consistent (Lee 2008). Although online calculators and guides help families predict their potential EFC, these guides are based on prior year Pell Grant schedules and the relationship between income and EFC is complicated and non-linear. In the years I examine, the maximum Pell Grant awards are set by amendments to the Higher Education Act. However, this legislation only specifies *authorized* annual maximum awards. The *appropriated* maximum award, which determines the actual Pell Grant schedule, is generally smaller than the authorized amount. Furthermore, the Department of Education releases the Pell Grant schedule after the end of calendar year, making it impossible for families to make real adjustments to most of the inputs used to determine EFC (e.g., adjusted gross income). Families might still misreport EFC inputs; however, many of these inputs are also reported to the IRS (e.g., adjusted gross income, number of dependents) and over one-third of all FAFSA applications are audited through the Department of Education's verification process.[18]

Nonetheless, I test for continuity and smoothness in the density of EFC to rule out the possibility that students perfectly manipulate their EFCs. Figure 1.6 displays the unconditional density of EFC, plotting the proportion of students in each $100 EFC interval. The x-axis measures the standardized distance from the Pell Grant eligibility cut-off. I limit the sample to

[18] The NPSAS contains an additional year of FAFSA information for continuing students who applied for federal aid for the following academic year. For these students, I test whether barely missing the eligibility threshold in the current year is correlated with any evidence of strategic behavior for the following year (e.g., bunching to the left of the new threshold). I find no evidence of this behavior (results available upon request).

students with $\tilde{EFC} \in [-2100, 2100]$ because of the large mass of individuals with an EFC of zero.[19] In 1996, a zero EFC corresponds to $\tilde{EFC} = -2140$, thus, this window prevents my estimates from being driven by the large increase in density at *EFC = 0*. Due to the smaller window, I use smaller bins ($100) than in other graphical analyses.

I use the McCrary (2008) test to determine whether the density of EFC is continuous across the threshold for Pell Grant eligibility. My method for testing continuity in the derivative around the cut-off is less precise, since there is presently no analog to the McCrary test statistic for the RK design. I follow Card et al. (2009), collapse the data into $100 EFC bins, and run the following regression:

$$N_b = \alpha + \beta \cdot \mathbf{1}[\tilde{EFC}_b < 0] + \gamma \cdot \mathbf{1}[\tilde{EFC}_b < 0] \cdot (\tilde{EFC}_b) + \sum_{\rho} [\lambda_\rho \cdot (\tilde{EFC}_b)^\rho] + \xi_b \tag{1.7}$$

Where b indexes bins, N_b is the number of students in bin b, ρ is a second order polynomial, $\tilde{EFC}_b$ is the distance from the eligibility threshold, and a test of $\gamma = 0$ estimates whether the density function is smooth. Figure 1.6 displays $\hat{\gamma}$ and the McCrary test statistic as well as the corresponding standard errors. I find no evidence that the level or the slope of the density changes discontinuously at the eligibility threshold.

I examine the distribution of predetermined student characteristics around the eligibility threshold, including race, gender, dependency status, average SAT score (first-year students only), and age; here bins represent $200 EFC intervals (Figures 1.7A through 1.7E). I also estimate equation (1.6), with up to a fifth degree polynomial in EFC, and test for discontinuous changes in the slope or level of baseline characteristic at the eligibility threshold. I display results

[19] In the years I examine, dependent students and independent students with dependents other than a spouse received an automatic zero EFC if (1) anyone in their household receive means tested benefits or their household was not required to file IRS Form 1040, and (2) their household income was less than $20,000.

from the specification using the optimal degree of polynomial (determined via the Akaike Information Criterion) for three different windows around the eligibility threshold in Table 1.3. For the majority of specifications, estimated coefficients are insignificant and all coefficients are quite small in magnitude.

Finally, I plot the density of EFC by institutional control and selectivity (Figures 1.8A through 1.8E). I find no evidence of a discontinuity in the density or its first derivative among public and nonselective private institutions. There is a positive increase in the density of EFC to the right of the Pell Grant eligibility threshold among selective nonprofit institutions, but the magnitude of the change in density is small and insignificant.[20]

1.5 Results

Figure 1.9 previews my main results. I pool observations from all schools across years and plot the relationship between Pell Grant aid, institutional aid, and standardized EFC (e.g., $200 indicates a student's EFC is $200 above the cut-off for Pell Grant eligibility). I collapse my data into $200 EFC bins and plot average institutional aid and average Pell aid by distance from the threshold for Pell Grant eligibility, where both institutional aid and Pell aid are residuals from a regression on year and institution fixed effects. Institutional aid is represented by hollow circles, with larger circles representing a greater number of students. Average Pell Grant aid is represented by the gray "X" markers. The black lines represent the linear fit of institutional aid on EFC, estimated separately on either side of the eligibility threshold and weighted by the number of students in the bin.[21] The dashed gray lines represent the 95 percent confidence

[20] Unfortunately, the NPSAS only contains information for accepted students who choose to enroll in a specific school, making it impossible to determine whether the density of EFC is smooth and continuous for all applicants.

[21] Figure 1.10 replicates this exercise, allowing for a more flexible fit of the relationship between institutional aid and EFC with a local linear regression. The resulting discontinuity and kink are very similar.

intervals for these estimates. Finally, the dashed black line represents the linear fit of average Pell Grant aid on EFC for eligible students. At the eligibility threshold, there is both discontinuous change in the relationship between EFC and institutional grant aid. For students who are ineligible for Pell Grants, there is a positive relationship between need and institutional aid, while for students who are eligible for Pell Grant aid, institutional aid is decreasing in need.

I replicate this exercise by sector (Figures 1.11A through 1.11C). Due to sample size constraints, I pool selective and nonselective public schools into a single category and likewise group nonselective nonprofit schools (parametric regression estimates, presented in the next section, suggest that schools within each of these groups respond similarly to need-based aid). The incidence of Pell Grant aid varies substantially between public and private schools, with public institutions appearing to supplement Pell Grants with increased institutional aid.[22] Private institutions' response to Pell Grant aid is more straightforward. There is a clear discontinuity in the slope of institutional aid to the left of the Pell Grant eligibility threshold and a negative, but insignificant change in the level of aid among nonselective private schools (Figure 1.11B). There is a small, insignificant jump in institutional aid for selective nonprofit schools, but the kink in the institutional aid schedule clearly dominates (Figure 1.11C).

1.5.1 Parametric RD and RK Estimates

Table 1.4 presents OLS and 2SLS estimates of equations (1.5) and (1.6) with a second degree polynomial in $\widetilde{EFC}$. The first two columns display the first stage and reduced form estimates, respectively. Columns 3 and 4 present separate RK and RD instrument variables estimates. Results are consistent with Figure 1.9 – RK estimates suggest that institutions capture

[22] To better illustrate the behavior of institutional aid around the eligibility threshold in the public sector, the left axis measures institutional aid while the right axis represents Pell aid.

around 22 cents of every Pell Grant dollar through a reduction in institutional aid while the RD estimator results in a point estimate of 0.32, suggesting schools *increase* institutional aid by over 30 cents for every dollar of Pell Grant aid. The test of equality of the RD and RK coefficients confirms that these estimated impact is statistically different. The test of equality also serves as a formal test of whether the impact of the Pell Grant Program on institutional pricing varies with EFC.

Before further investigating the surprising result suggested by the RD estimator – that schools respond to each additional dollar of Pell Grant aid by increasing institutional aid – I test how robust my main results are to difference specifications by varying the window and polynomial in $E\tilde{F}C$ to confirm that this result is not an artifact of a particular specification (Table 1.5). I use three windows of standardized ECF: $E\tilde{F}C < 10{,}000$, $E\tilde{F}C \in [-4000, 4000]$, and $E\tilde{F}C \in [-3000, 3000]$.[23] For each window, I include up to a third degree polynomial in standardized EFC and use the Aikake Information Criterion (AIC) to determine the optimal degree of polynomial. For all but the largest window, a linear term in standardized EFC provides the best fit to the data. Results are consistent across windows and polynomials in $E\tilde{F}C$.

1.6 A Framework for Understanding Differences in RK and RD Estimates

Would a profit-maximizing firm ever pass-through more than 100 percent of a subsidy to consumers? When firms have market power, the economic incidence of a tax or subsidy may exceed 100 percent, but a simple model suggests that my result would not occur without very specific patterns of student demand or a departure from profit-maximization.

[23] The largest window encompasses students with an AGI ranging from $0 to approximately $90,000, the second window includes families whose AGI falls between $20,000 and $60,000, and the smallest window restricts the analysis to families with an AGI between $25,000 and $50,000.

First, suppose a profit-maximizing monopolist serving N distinct student groups solves:

$$\max_{p_1,\ldots,p_N} \pi = \sum_{i=1}^{N} Q_i(p_i)(p_i - c)$$

where Q_i is the demand of students in group i and c is the marginal cost of serving an additional student. For simplicity, I assume marginal costs are constant, both in the number of students served and across student groups, which is reasonable if instruction and facilities make up the majority of expenses. The school charges students in group i a price that is equal to overall tuition (which does not vary across groups) minus institutional aid (which may vary across groups). Groups are defined by students observable characteristics (e.g., demographic characteristics, EFC), and schools use these characteristics to practice price discrimination. This is a static problem, where a school's behavior in the current period does not affect cost or demand in future periods.

A profit-maximizing monopolist charges group i students price $p_i = c\mu_i$, where $\mu_i = \left(\frac{e_i}{e_i + 1}\right)$ and e_i is the price elasticity of demand for students in group i. When federal need-based aid (s) is introduced, the school charges $p_i = (c - s)\mu_i$, where $s < c \ \forall i$. The change in the final price paid by students in group i in response to the subsidy will be:

$$\frac{dp_i}{ds} = -\mu_i + (c - s)\frac{d\mu_i}{ds} \tag{1.8}$$

For instance, $\frac{dp_i}{ds} = 0$ indicates that the school fully captures every additional dollar of the subsidy, while $\frac{dp_i}{ds} = -1$ indicates subsidies are fully passed-through to students. The sign of $\frac{dp_i}{ds}$ depends on both the elasticity and the curvature of the demand function for students in group i

(Bulow and Pfleiderer 1983; Weyl and Fabinger 2011). If demand is log-concave, $\frac{dp_i}{ds} > -1$, and schools capture a portion of students' Pell Grant aid by increasing prices (decreasing institutional aid).[24] If demand is log-convex, $\frac{dp_i}{ds} < -1$, and schools respond to Pell Grant aid by decreasing prices (increasing institutional aid), the result suggested by the RD estimator.[25]

However, the increase in institutional aid combined with the change in the slope of the institutional aid-EFC schedule at the threshold, with institutional transfers decreasing with every additional dollar of Pell Grant aid, is more surprising. If student demand is log-convex, then institutional transfers should increase as Pell Grant aid increases. There would have to be sharp changes in the demand functions of students near the eligibility threshold to account for the patterns of institutional aid provision I observe. Specifically, the initial $400 Pell Grant award would have to move students from a log-concave portion of the demand curve to a log-convex portion, requiring the existence of an inflection point. This is unlikely, since the eligibility threshold for Pell Grant aid changes over time, while pricing patterns are persistent (results available upon request).

Conversely, suppose a subset of schools have a different objective function, and maximize weighted student enrollment, where weights vary across student groups:

---

[24] The price set by a school has two components: tuition and institutional aid: $p_i = t - a_i$. Since schools set tuition before Pell Grant awards are announced, only institutional aid responds to Pell Grant awards, thus $\frac{dp_i}{ds} = -\frac{da_i}{ds}$.

[25] In the short-run, this model can be easily generalized to represent institutional pricing with monopolistically competitive firms offering differentiated products. In this case, student demand will depend not only on an institution's price but the prices offered by competitors, $Q_i = Q_i\left(p_i, p_{j \neq i}\right)$, and pricing will also depend on the cross-price elasticities of demand. Pass-through will be decreasing in the number of competitors in the market and the degree of substitutability between programs offered by institutions. In the long-run, incidence will depend on the ease of entry into a specific market. Additionally, a substantial minority of institutions are monopolists. In 2009, 17 percent of all institutions eligible to disburse federal aid were the only institution in their county.

$$\max_{p_1,\ldots,p_N} W = \sum_{i=1}^{N} \alpha_i Q_i(p_i) \qquad \text{s.t.} \qquad \sum_{i=1}^{N} Q_i(p_i)(p_i - c) \geq 0$$

The constraint stems from the requirement that in a static model, expenditures cannot exceed revenue. If the constraint is binding, schools will offer a schedule of prices that vary by demand elasticity as well as the weight placed on the group in the schools objective function (α_i) and the marginal "utility" of revenue (represented by the Lagrange multiplier): $p_i = (c - \tilde{\alpha}_i)\mu_i$, where $\tilde{\alpha}_i$ is the weight on students in group *i* divided by the Lagrange multiplier. If being labeled as a Pell Grant recipient affects this weight, schools' pricing response to subsidy *s* is now:

(1.9) $$\frac{dp_i}{ds} = -\left(\frac{d\tilde{\alpha}_i}{ds} + 1\right)\mu_i + \left(c - \tilde{\alpha}_i(s) - s\right)\frac{d\mu_i}{ds}$$

Equation (1.9) implies that if Pell Grant recipients receive a positive weight in the school's objective function (e.g., $\tilde{\alpha}_i(s) > 0$), the second term will be smaller than in the case of static profit maximization. Furthermore, if Pell Grant recipients' weights are larger than those of observationally similar students who do not qualify for Pell Grant aid (e.g., $\frac{d\tilde{\alpha}_i}{ds} > 0$), the first term will be larger. If either of these terms is positive, these schools will capture a smaller portion of Pell Grant aid. Furthermore, rearranging equation (1.9):

(1.10) $$\frac{dp_i}{ds} = \left\{-\mu_i + (c - s)\frac{d\mu_i}{ds}\right\} - \left\{\mu_i \frac{d\tilde{\alpha}_i}{ds} + \tilde{\alpha}_i(s)\frac{d\mu_i}{ds}\right\}$$

Here the first term represents the pass-through of outside student aid due to profit maximization (or cost minimization), while the second term accounts for a school's willingness to pay for Pell Grant recipients. If, in the neighborhood of the cut-off for Pell Grant eligibility, $\frac{d\tilde{\alpha}_i}{ds}$ does not vary with s for Pell Grant recipients (e.g., if being a Pell Grant recipient increases your weight in the school's objective function by a constant amount), the relationship between the prices for group i students and Pell Grant aid can be approximated by: $p_i \approx \tau_0 \mathbf{1}[s_i > 0] + \tau_1 s_i + u_i$. Here, p_i is the final price faced by students in group i, τ_0 and τ_1 represent willingness to pay for Pell Grant recipients and pass-through of each additional dollar of Pell Grant aid, respectively, and u_i is an idiosyncratic error term.[26]

There are a number of reasons why schools might treat Pell Grant recipients differently than other students. First, schools might have objectives beyond profit maximization, such as increasing school-wide diversity or maximizing (weighted) student welfare. Schools might solve a dynamic problem where additional Pell Grant recipients in the current period increase the expected value of the stream of future revenue (or reduce the expected value of the stream of future costs). For example, schools that serve a larger number of Pell Grant recipients might receive more funding from state legislatures in the long-run or experience an increase in student demand. For instance, in recent years, the U.S. News and World Report began incorporating a measure of Pell Grant receipt in its school ranking calculations. For the purposes of this paper, I remain agnostic as to the reasons schools might treat Pell Grant recipients differently from students who barely miss the cut-off for eligibility.

1.6.1 Estimating the Multiple Treatment Dimensions of Pell Grant Receipt

Equation (1.10) suggests that the "treatment" of receiving a Pell Grant affects prices along two dimensions: a school's willingness to pay for Pell Grant recipients (τ_0) and ability to appropriate outside aid due to the pass-through of cost decreases (τ_1). To see how these two

---

[26] This approximation also assumes that in the neighborhood of the Pell Grant eligibility threshold, each additional dollar of Pell Grant aid does not lead to large changes in the log-curvature of demand.

dimensions are related to RD and RK estimates, consider a simplified version of equation (1.6), the reduced form impact of Pell Grant eligibility on institutional aid in a specific school and year:

$$y_i = \beta \cdot \mathbf{1}\left[\tilde{EFC}_i < 0\right] + \gamma \cdot \mathbf{1}\left[\tilde{EFC}_i < 0\right] \cdot \left(\tilde{EFC}_i\right) + \pi\left(\tilde{EFC}_i\right) + \varepsilon_i$$

Furthermore, assume that all eligible students receive a Pell Grant, where the minimum award is \$400 (e.g., "sharp" RD/RK).

The RD design provides a reduced form estimate of the "treatment" of Pell Grant receipt, where $\beta = \tau_0 + \tau_1 \cdot 400$ and $\tau_{RD} = \frac{\tau_0}{400} + \tau_1$, which confounds the school's ability to capture an additional dollar of outside aid with its willingness to pay for students labeled as Pell Grant recipients (see Appendix). When these two dimensions have opposite signs, RD estimates will not identify the magnitude and sign of either dimension.

The RK design will consistently estimate the pass-through of an additional dollar of outside aid, under the assumption that τ_1 is constant in the neighborhood of the cut-off for Pell Grant eligibility (see Appendix). Since $\tau_{RK} = \tau_1$ and the RK/RD design is fuzzy:

$$\begin{aligned} \hat{\tau}_1 &= \hat{\tau}_{RK} \\ \hat{\tau}_0 &= \left(\hat{\tau}_{RD} - \hat{\tau}_{RK}\right) \cdot Pell\left(efc_0\right) \end{aligned} \tag{1.11}$$

Where $Pell\left(efc_0\right)$ is the minimum Pell Grant award, $\hat{\tau}_{RD}$ and $\hat{\tau}_{RK}$ are the RD and RK estimators, respectively, $\hat{\tau}_0$ is the estimated willingness to pay for Pell Grant recipients, and $\hat{\tau}_1$ is the pass-through of Pell Grant aid from the student to the school. The appendix provides further details for the derivation of these parameters.

Table 1.6 presents estimates of the capture and willingness to pay parameters for the pooled sample (Panel A) and by sector (Panel B). I use the delta method to calculate standard errors. Across all institutions, estimated pass-through is 0.22, suggesting institutions receive 22 cents of every additional dollar of Pell Grant aid. However, due to schools' willingness to pay for Pell Grant recipients, Pell Grant recipients experience a $260 increase in institutional aid. Since students ineligible for Pell Grants received $1,800 in institutional aid on average (including students that did not receive any institutional aid), this transfer represents a 14 percent increase in the expected value of institutional aid.[27] However, only Pell Grant recipients near the eligibility threshold benefit from these transfers, and these students make up less than a third of all recipients. For the remainder of Pell Grant recipients, schools' ability to capture Pell Grant aid outweighs willingness to pay for needy students.

Figures 1.10A through 1.10C suggest that pass-through Pell Grant aid and willingness to pay for Pell Grant recipients vary across sectors. I test for differences in behavior by fully interacting $Pell_{it}$ with a vector of indicators for the different sectors of higher education. Private institutions do not demonstrate a willingness to pay for Pell Grant recipients and 13 to 15 cents of every Pell Grant dollar is passed-through from students to nonselective institutions. Conversely, public schools increase institutional aid for recipients by $300 to $600 in public schools. The difference in willingness to pay between selective and nonselective public schools is only marginally significant. This additional aid represents a 140 percent increase in the expected value of institutional grants among nonselective public school students and a 90 percent increase for selective public school students.[28]

While public schools appropriate 17 to 18 cents of every Pell Grant dollar, pass-through of Pell Grant aid is the largest among selective nonprofit institutions. These schools capture 69

---

[27] This calculation includes students within $10,000 of the Pell Grant eligibility threshold. However, if I limit the distance to be $4,000, estimated institutional aid is quite similar ($1850).

[28] On average, Pell ineligible students receive approximately $230 in institutional aid from nonselective public schools and $700 from selective public schools.

cents every Pell Grant dollar, while any willingness to pay for Pell Grant recipients is quickly overtaken. This result suggests that selective nonprofits either serve students with less elastic demand or have greater market power.

1.6.2 Heterogeneity by Student and Market Characteristics

To determine whether differences in student demand can explain differences in pass-through between sectors, I examine heterogeneity in pass-through and schools willingness to pay for Pell Grant recipients across three student demographic groups, defined by race (white versus nonwhite), dependency status, and gender (Table 1.7). If students with similar characteristics have relatively similar demand elasticities, this analysis provides a test of whether the greater degree of pass-through in the selective nonprofit sector stems from differences in the demand of the students these schools serve. Specifically, if selective nonprofits serve a segment of the market is with less elastic demand, pass-through will be greater without any differences in these schools' underlying objectives.

I find that across all demographic groups, pass-through of Pell Grant aid in the selective nonprofit sector is significantly greater than in other sectors, except in the case of independent students.[29] Additionally, public schools display a willingness to pay for Pell Grant recipients across all groups. These results suggest that differences in the characteristics of students served cannot fully explain selective nonprofit institutions' large degree of capture or public schools' valuation of students receiving Pell Grant aid.

Measuring schools' market power is a more difficult task. I measure the *ex ante* degree of competition in a school's market using a Herfindahl index of institutional shares of the

[29] In independent students in all sectors experience the smallest degree of crowd-out, suggesting these students have more elastic demand than dependent students.

undergraduate population during the prior academic year. I define the market served by a particular institution to be the county in which it is located, since the median distance a student travels to attend a nonselective institution is 15 miles (Horn and Nevill 2006), and use data from the IPEDS to measure the total number of undergraduate full-time equivalent (FTE) students in a county and institutional shares for NPSAS and non-NPSAS schools.[30] Although some selective schools effectively serve a national market, I find evidence that Pell Grant receipt causes some students to switch from attending nonselective schools to selective institutions, suggesting students may be evaluating their choices in their local market.

To test whether pass-through varies by market structure, I create a dichotomous measure of concentration based on the index (H>0.25) and estimate equation (1.11), fully interacting $Pell_{it}$, $\mathbf{1}\left[E\tilde{F}C_{it} < 0\right]$, and $\mathbf{1}\left[E\tilde{F}C_{it} < 0\right]\left(E\tilde{F}C_{i}\right)$ with this measure. Table 1.8 displays these results (estimates of willingness to pay by market concentration are available upon request). In column 2, I consider all other institutions when constructing the index; in column 3, I only consider institutions with similar selectivity (e.g., assuming that selective nonprofit schools do not compete with for-profit schools for students). I find some evidence that selective nonprofit institutions respond to competitors – in markets with few similarly selective schools, these institutions capture 79 cents of every Pell Grant dollar, while in more competitive markets, only 44 cents of every Pell dollar are passed through. However, my measure of market power is blunt and does not account for endogenous entry decisions.

My results represent the short-run incidence of Pell Grant aid. In the long-run, increases in competition may limit schools' ability to capture student aid. Although the supply of public

[30] Unfortunately, prior to 2000, the IPEDS data files do not accurately represent the presence of for-profit institutions among the set of schools eligible to disburse federal student aid. However, as shown in column 1 of Table 1.9, estimates of pass-through are quite similar for this truncated sample.

institutions is relatively fixed, Cellini (2010) shows that student aid increases lead to for-profit entry. If for-profit institutions retain captured Pell Grant aid as profits, my results provide a rationale for this response. An increase in number of schools should reduce the ability of schools to capture Pell Grant aid and in the long-run, institutional capture should be driven to zero. Incidence analysis in this case is complicated by the fact that captured Pell Grant funds in the present period ultimately lead to an expansion provision of higher education. Although current Pell recipients lose out, new students, who would not have otherwise attended college, will gain from the ability of schools to capture Pell aid. However, the market for higher education also has substantial barriers to entry, since schools face large fixed costs (e.g., investments in facilities). Schools also must gain accreditation and demonstrate a sufficiently high level of enrollment for two years before their students are eligible for Pell Grant aid.

1.6.3 Evaluating Alternate Explanations for Institutional Pricing

Thus far, I have attributed differences in institutional pricing responses to Pell Grant aid to differences in institutional objectives and market power. However, there are other potential explanations for this behavior. First, differences in unmet need and institutional policies across sectors could potentially explain differences in estimated crowd-out. For instance, since public schools charge significantly lower prices than private institutions, institutional aid may mechanically fall if increases in Pell Grant aid reduce students' unmet need to zero. State need-based aid may be distributed differently across sectors, also contributing to this effect. Figures 1.12 and 1.13 explore this possibility and plot the percentage of students with any unmet need and average unmet need and by EFC and sector, where unmet need is defined as the difference between a student's cost of attendance and her expected family contribution, Pell Grant and other federal grant aid, and state grant aid. Across all sectors, over 90 percent of students near the Pell

Grant eligibility threshold had remaining unmet need, and on average, these students faced an additional $10,000 in expected education-related expenses that are not covered by federal or state grant aid. Even students attending nonselective public institutions –schools with the lowest cost of attendance – had substantial remaining need.

Second, students may respond to Pell Grant generosity by upgrading to a higher quality institution. In this case, price increases would be expected, as students are receiving a more valuable product. Although I find some evidence of sorting across sectors at the eligibility threshold – with a small, discrete increase in the probability of attending a selective institution as the eligibility threshold is crossed – there is no evidence of a kink. Since selectivity is just one dimension of school quality, I also test for evidence of quality upgrading by examining institutional revenue, expenditures, and the outcomes of former students. I use information from the IPEDS linked to NPSAS institutions to create measures of revenue and expenditures, including tuition and total revenue per full-time equivalent (FTE) student and institutional grants, instruction-related expenditures, and expenditures on student services per FTE.[31] Finally, I use the Department of Education's official cohort default rate, which measures the proportion of individuals defaulting on their federal loans within the two years, to measure the outcomes of former students.

I little evidence of upgrading along any of these measures of school quality and in many cases, find evidence of a negative relationship between Pell Grant generosity and school quality (Table 1.9). I estimate a positive relationship between Pell Grant aid and expenditures on institutional grants and instruction for students attending public schools, but the magnitudes of

[31] I use prior-year revenue and expenditure data to create these measures. Unfortunately, the IPEDS only began collected revenue and expenditure data for the majority of schools in 2000, thus, when examining these measures of quality, my sample is limited to students attending institutions in 2004 and 2008.

these effects are quite small. A $1000 increase in Pell Grant aid is correlated with a $14 (1 percent) increase in institutional grants/FTE offered by nonselective public schools. For selective public schools, a $1000 increase in Pell Grant aid is correlated with a $22 (1 percent) increase in institutional aid and a $132 (2 percent) increase in instruction-related expenditures for students attending selective public schools. Pell Grant aid is negatively correlated with student loan default rates in the for-profit sector, where a $1000 increase in Pell Grant aid is correlated with a 0.9 percentage point reduction in the default rate of graduating students. However, among students attending selective nonprofit institutions – the sector which shows the highest degree of crowd-out – there is no evidence of quality upgrading.

1.7 Incidence Across All Pell Grant Recipients

So far, I have only focused on estimating the incidence of Pell Grant aid in the neighborhood of the cut-off for Pell Grant eligibility. With stronger assumptions, I can use the observable relationship between institutional aid and EFC for ineligible students to estimate the incidence of the Pell Grant program across all students. Specifically, I assume that the relationship between institutional aid and EFC for ineligible students provides a valid counterfactual for what the relationship between institutional aid and EFC would have been for Pell Grant recipients in the absence of the Pell Grant Program. For this approach to work, heterogeneous treatment effects must be linear. Specifically, the pass-through of Pell Grant aid and schools' willingness to pay for Pell Grant recipients must be constant in the amount of Pell Grant aid

Figure 1.14 provides an illustration of my approach. The shaded area under the Pell Grant curve represents the total amount of aid directed towards Pell Grant recipients by the federal government. The solid lines represent the observed relationship between institutional aid and EFC for eligible and ineligible students, while the dashed line represents counterfactual institutional aid for eligible students in the absence of the Pell Grant Program. In other words, each point along this line represents the amount of institutional aid a student with a particular EFC would have received had the Pell Grant Program not existed. The difference between the area under the first curve (counterfactual institutional aid) and the second curve (actual institutional aid) represents institutional capture ($A - B$). The ratio of total capture to total Pell aid, $\frac{A-B}{Total\ Pell}$, represents the percentage of Pell Grant aid captured by institutions, and is also the average treatment effect of Pell Grants on institutional aid.

Across all sectors, every dollar of Pell Grant aid reduces students' effective prices by 84 cents, with institutions appropriating the remaining 16 cents through price discrimination (Table 1.10). Nonselective private institutions, a category encompassing nonprofit and for-profit schools, receive 18 cents of every Pell Grant dollar while selective nonprofit institutions capture 79 cents. In the public sector, net crowd-out of Pell Grant aid is close to zero; the point estimate is small and only marginally insignificant. However, this result masks important heterogeneity – transfers to students close to the eligibility threshold are offset by decreases in institutional aid for the neediest Pell Grant recipients (Figure 1.10A).

1.8 Conclusions

Although low-income students are the statutory recipients of Pell Grant aid, they do not receive the full benefit of these subsidies. Using a combined regression discontinuity and regression kink approach, I estimate the impact of Pell Grants on institutional aid. I show that schools strategically respond to changes in federal need-based aid by systematically altering the distribution of institutional aid. Overall, I estimate that institutions capture 16 percent of all Pell

Grant aid. However, this result masks important variation in pass-through across sectors and across students with different levels of need.

RK and RD designs yield conflicting estimates of the impact of Pell Grant aid on college pricing, with RK estimates suggesting schools capture Pell Grant aid and the RD estimator implying schools supplement Pell Grants with *increased* institutional aid. I show that these disparate estimates can be reconciled using a framework in which schools place different weights on students with different characteristics. In this case, the "treatment" of Pell Grant aid has two dimensions: the additional dollar of outside aid that the school would like to capture and the school's willingness to pay for Pell Grant recipients.

Through a combined RD/RK approach, I separately identify schools' willingness to pay for students categorized as needy and the pricing response to outside subsidies. The RD design only identifies the reduced form impact of these two dimensions, and for RD estimates, schools' willingness to pay dominates their ability to capture outside aid. Using the combined RD/RK approach, I estimate that less than one third of Pell Grant recipients benefit from these transfers, since schools' ability to capture Pell Grant aid quickly overtakes their willingness to pay for needy students. My paper is the first to combine RD and RK estimators to distinguish between different treatment dimensions.

The Pell Grant Program provides an especially stark example of how a multidimensional treatment affects RD estimates. However, in other circumstances where both a discontinuity and a kink are present, my results suggest that additional information is present in the kink, and this information may provide insight into the channels through which the "treatment" of interest works. In a number of the studies cited by Lee and Lemieux's (2010) survey on the RD design that examine the impact of a continuous endogenous regressor, the deterministic relationship between the endogenous regressor and assignment variable leads to both a discontinuity and a kink. For instance, in cases where a minimum class size rule leads to a discontinuous relationship between total enrollment and class size (e.g., Angrist and Lavy 1999; Hoxby 2000; Urquiola 2006), this rule creates both a discontinuity and a kink.[32] If, for instance, the creation of an additional classroom leads to smaller classes *and* sorting of children by achievement, behavior, or some other dimension (e.g., Lazear 2001), the discontinuity and the kink could potentially be used to separately analyze the influence of these dimensions on educational outcomes.

My paper also provides insight into the industrial organization of higher education. I show how schools' responses to Pell Grant aid illustrate differences in schools' objectives and market power across sectors. Public schools demonstrate a positive willingness to pay for Pell Grant recipients. Overall, selective nonprofit institutions capture close to 80 percent of their students' Pell Grants. Across different student demographic groups, I estimate a similar degree of capture students attending selective nonprofits, suggesting these schools' extensive ability to appropriate Pell Grant aid stems from a greater degree of market power rather than differences in student demand. Although the net crowd-out of Pell Grants in the public sector is close to zero, increases in institutional aid for recipients near the eligibility threshold come at the expense of the neediest Pell recipients.

Finally, I find no evidence that for-profit institutions behave differently than other nonselective schools in the private sector in their response to Pell Grant aid, and combined, schools in this sector capture just 17 percent of their students' Pell Grant aid. However, in many

[32] For example, if the rule mandates a maximum class size of $\overline{N}$, when enrollment reaches $\overline{N}+1$, average class size changes discontinuously from $\overline{N}$ to $\frac{\overline{N}+1}{2}$. This rule also leads to a kink in the relationship between average class size and total enrollment. When enrollment is less than $\overline{N}$, the slope of relationship between class size and total enrollment is 1. When class size is greater than $\overline{N}$, but less than $2\overline{N}$, the slope of the relationship between class size and total enrollment is 0.5.

for-profit institutions, the majority of students receive Pell Grants. It may be easiest for these institutions to benefit from Pell Grant generosity by raising the list price of tuition. Consistent with this view, Cellini and Goldin (2012) show that in the for-profit sector, schools eligible to distribute federal student aid charge a list price that is 75 percent higher than ineligible schools with similar characteristics.

Under the stronger assumption that the distribution of institutional aid to ineligible students near the threshold provides a valid counterfactual for the distribution of institutional aid in the absence of the Pell Grant Program, I calculate that schools capture 16 percent of all Pell Grant aid. In 2011, the federal government distributed $35 billion in Pell Grants to 9.5 million students. My results suggest that institutions captured $6 billion of this aid.

1.9 Appendix: Regression Discontinuity Estimation with a Multidimensional Treatment

In this appendix, I provide a general example of how a multidimensional treatment will affect regression discontinuity (RD) design estimates. Additionally, I show how using a regression kink (RK) design, in combination with a RD design, allows estimation of more than one treatment dimension. Finally, I show how this approach is applied in the case of the Pell Grant Program.

Let Y be the outcome of interest, where $Y = y(T, X, U)$ and T is the "treatment" of interest and is continuous and potentially endogenous. X and U are covariates, where X is observable, U is unobservable, and both are determined prior to T. Finally, T is a deterministic function of X, $T = t(X)$, and the data generating processes for Y and T are:

(B1) $$Y = f(T, \tau) + g(X) + U$$

(B2) $$T = \beta_0 1[X \leq x_0] + \beta_1 X \cdot 1[X \leq x_0] + h(X)$$

Where $h(X)$ is continuously differentiable in the neighborhood of x_0. In this case, the deterministic relationship between T and X leads to both a change in the level and in the first derivative of T at x_0. I assume that treatment effects do not vary with X or U, but this assumption could be relaxed without affecting my main conclusions. Finally, $F_U(u)$ is the cdf of U and $F_{X|U}(x \mid u)$ is the conditional cdf of X.

Under the following identifying assumptions, the RD estimator will approximate random assignment (Hahn et al. 2001; Lee and Lemieux 2010):

RD1 (Regularity): $y(t, x, u)$ is continuous in x in the neighborhood of x_0 and $f_U(x_0) > 0$.

RD2 (First Stage): T is a known function, continuous on $\left(-\infty, x_0\right)$ and $\left(x_0, \infty\right)$, but

$\lim_{\varepsilon \uparrow 0} E\left[T \mid X = x_0 + \varepsilon\right] \neq \lim_{\varepsilon \downarrow 0} E\left[T \mid X = x_0 + \varepsilon\right]$.

RD3 (Continuous conditional density of the assignment variable): $f_{X|U}\left(x \mid u\right)$ is continuous in x in the neighborhood of x_0 for every u. This condition means that observations have imperfect control over X and rules out sorting in response to the treatment.

Consider two different forms of $f\left(T, \tau\right)$:

(B3) $$f\left(T, \tau\right) = \tau_1 T$$

(B4) $$f\left(T, \tau\right) = \tau_0 1[T > 0] + \tau_1 T$$

If equation (B3) describes $f\left(T, \tau\right)$, "treatment" has a single dimension, as is generally assumed in RD designs, the RD estimator equals:

$$\tau_{RD} = \frac{\lim_{\varepsilon \uparrow 0} E\left[Y \mid X = x_0 + \varepsilon\right] - \lim_{\varepsilon \downarrow 0} E\left[Y \mid X = x_0 + \varepsilon\right]}{\lim_{\varepsilon \uparrow 0} E\left[T \mid X = x_0 + \varepsilon\right] - \lim_{\varepsilon \downarrow 0} E\left[T \mid X = x_0 + \varepsilon\right]} = \tau_1$$

If instead, T is multidimensional and equation (B4) describes $f\left(T, \tau\right)$, the RD estimator equals:

$$\tau_{RD} = \tau_1 + \frac{\tau_0}{T\left(x_0\right)}.$$

To see this, note that the numerator of the RD estimator is equal to:

$$\lim_{\varepsilon \uparrow 0} E\left[\tau_0 1[T > 0] + \tau_1 T + g\left(X\right) + U \mid X = x_0 + \varepsilon\right] - \lim_{\varepsilon \downarrow 0} E\left[\tau_0 1[T > 0] + \tau_1 T + g\left(X\right) + U \mid X = x_0 + \varepsilon\right]$$

Assumptions RD1 and RD3 imply: $\lim_{\varepsilon\uparrow 0} E[g(X)+U \mid X = x_0 + \varepsilon] = \lim_{\varepsilon\downarrow 0} E[g(X)+U \mid X = x_0 + \varepsilon]$.

Since $\lim_{\varepsilon\uparrow 0} E[h(X) \mid X = x_0 + \varepsilon] = \lim_{\varepsilon\downarrow 0} E[h(X) \mid X = x_0 + \varepsilon]$ by assumption, the RD numerator is equal to $\tau_0 + \tau_1(\beta_0 + \beta_1 x_0)$ and the RD estimator equals:

(B5) $$\tau_{RD} = \tau_1 + \frac{\tau_0}{\beta_0 + \beta_1 x_0} = \tau_1 + \frac{\tau_0}{T(x_0)}$$

Thus, when the treatment has more than one dimension, the RD estimator only recovers the reduced form impact of these dimensions. In this case, with the RD design alone, it is not possible to separately identify τ_0 and τ_1. However, since the deterministic relationship between T and X leads to both a discontinuous change in the level and a discontinuous change in the slope of T at x_0, it is possible to separately identify these dimensions using a combined RD and RK approach.

In addition to the RD identifying assumptions, the RK design requires the following additional assumptions (Card et al., 2009):

RK1 (Regularity): $\frac{\partial y(t,x,u)}{\partial x}$ is continuous in x in the neighborhood of x_0.[33]

RK2 (First Stage): T is continuously differentiable on $(-\infty, x_0)$ and (x_0, ∞), but

$$\lim_{\varepsilon\uparrow 0} \frac{\partial E[T \mid X = x_0 + \varepsilon]}{\partial x} \neq \lim_{\varepsilon\downarrow 0} \frac{\partial E[T \mid X = x_0 + \varepsilon]}{\partial x}.$$

[33] Card et al. (2009) include the additional assumption that $\frac{\partial y(t,x,u)}{\partial t}$ is continuous in t. If treatment is multidimensional, this condition is violated. Comparisons of RD and RK estimators allows for a test of whether this condition is met.

RD3 (Continuously differentiable conditional density of the assignment variable): $f_{X|U}(x \mid u)$ is continuously differentiable in *x* in the neighborhood of x_0 for every *u.*

If these conditions are met, regardless of whether $f(T,\tau)$ is represented by equation (B3) or (B4), the RK estimator equals:

$$\tau_{RK} = \frac{\lim_{\varepsilon \uparrow 0} \frac{\partial E[Y \mid X = x_0 + \varepsilon]}{\partial x} - \lim_{\varepsilon \downarrow 0} \frac{\partial E[Y \mid X = x_0 + \varepsilon]}{\partial x}}{\lim_{\varepsilon \uparrow 0} \frac{\partial E[T \mid X = x_0 + \varepsilon]}{\partial x} - \lim_{\varepsilon \downarrow 0} \frac{\partial E[T \mid X = x_0 + \varepsilon]}{\partial x}} = \tau_1$$

To see this, first note that the numerator equals:

$$\lim_{\varepsilon \uparrow 0} \frac{\partial E[\tau_0 1[T > 0] + \tau_1 T + g(X) + U \mid X = x_0 + \varepsilon \mid X = x_0 + \varepsilon]}{\partial x} - \lim_{\varepsilon \downarrow 0} \frac{\partial E[\tau_0 1[T > 0] + \tau_1 T + g(X) + U \mid X = x_0 + \varepsilon \mid X = x_0 + \varepsilon]}{\partial x}$$

By assumptions RK1 and RK3, $\lim_{\varepsilon \uparrow 0} \frac{\partial E[g(X) + U \mid X = x_0 + \varepsilon]}{\partial x} = \lim_{\varepsilon \downarrow 0} \frac{\partial E[g(X) + U \mid X = x_0 + \varepsilon]}{\partial x}$,

$\lim_{\varepsilon \uparrow 0} \frac{\partial E[1[T > 0] \mid X = x_0 + \varepsilon]}{\partial x} = \lim_{\varepsilon \downarrow 0} \frac{\partial E[1[T > 0] \mid X = x_0 + \varepsilon]}{\partial x} = 0$, regardless of the value of τ_0, and

$\lim_{\varepsilon \uparrow 0} \frac{\partial E[h(X) \mid X = x_0 + \varepsilon]}{\partial x} = \lim_{\varepsilon \downarrow 0} \frac{\partial E[h(X) \mid X = x_0 + \varepsilon]}{\partial x}$, by assumption. Thus, the RK numerator equals $\tau_1 \left(\lim_{\varepsilon \uparrow 0} \frac{\partial E[T \mid X = x_0 + \varepsilon]}{\partial x} - \lim_{\varepsilon \downarrow 0} \frac{\partial E[T \mid X = x_0 + \varepsilon]}{\partial x} \right)$ and the RK estimator equals:

(B6) $$\tau_{RK} = \tau_1$$

Furthermore, if the treatment has two dimensions, as described in equation (B4), the RD and RK estimators can be combined to identify both τ_0 and τ_1. The RK estimator identifies τ_1, and combining (B5) and (B6) allows for identification of the second treatment dimension:

(B7) $$\tau_0 = (\tau_{RD} - \tau_{RK})T(x_0)$$

If $f(T,\tau)$ has higher order terms, then $\tau_{RD} = \dfrac{\tau_0}{T(x_0)} + \tau_1 + \tau_2 T(x_0) + ... + \tau_p T(x_0)^{p-1}$ and $\tau_{RK} = \tau_1 + \tau_2 T(x_0) + ... + \tau_p T(x_0)^{p-1}$ where p is the order of the polynomial in T. Thus, using a combined RD/RK approach, it is always possible to identify τ_0, or the discrete change in the outcome that occurs when $T > 0$, but it is not possible to separately recover higher order terms without discontinuities in higher order derivatives of T.

B.1 Identification of multiple treatment dimensions in the case of the Pell Grant Program

In the case of the Pell Grant Program, $Y = y(Pell, EFC, U)$ represents institutional aid. Since not every student submits an application for federal aid, Pell Grant aid is not completely determined by a student's EFC, and the RD/RK designs will be fuzzy. The data generating processes for Y and $Pell$ are:

(B8) $$Y = f(Pell, \tau) + g(EFC) + U$$

(B9) $$Pell = \pi \cdot \mathbf{1}[EFC < efc_0](400 - (EFC - efc_0))$$

Where efc_0 is the cut-off for Pell Grant eligibility, and $\pi \in \{0,1\}$ (e.g., the probability a student applies for federal aid) is a random variable where $E[\pi] > 0 \, \forall efc$. Although π may also depend on *EFC*, since the decision to apply is determined prior to an individual receives their Pell Grant award, I assume that $\pi = \pi(EFC)$ is continuous and smooth in the neighborhood of efc_0.

My model suggests that Pell Grant aid may affect institutional aid provision through two dimensions: by altering a school's willingness to pay (τ_0) and through schools' ability to capture outside aid due to the pass-through of demand increases (τ_1):

(B10) $$f(Pell,\tau) = \tau_0 \mathbf{1}[Pell > 0] + \tau_1 Pell$$

The numerator of the RD estimator will be equal to:

$$\lim_{\varepsilon \uparrow 0} E[f(Pell,\tau) + g(EFC) + U \mid EFC = efc_0 + \varepsilon] - \lim_{\varepsilon \downarrow 0} E[f(Pell,\tau) + g(EFC) + U \mid EFC = efc_0 + \varepsilon]$$

Since $\lim_{\varepsilon \uparrow 0} E[g(EFC) + U \mid EFC = efc_0 + \varepsilon] = \lim_{\varepsilon \downarrow 0} E[g(EFC) + U \mid EFC = efc_0 + \varepsilon]$,

by assumptions RD1 and RD3, the RD numerator is equal to:

$$\lim_{\varepsilon \uparrow 0} E[\tau_0 \mathbf{1}[Pell > 0] + \tau_1 Pell \mid EFC = efc_0 + \varepsilon] - \lim_{\varepsilon \downarrow 0} E[\tau_0 \mathbf{1}[Pell > 0] + \tau_1 Pell \mid EFC = efc_0 + \varepsilon]$$

$$\begin{aligned} &= \tau_0 \left(\lim_{\varepsilon \uparrow 0} E[\mathbf{1}[Pell > 0] \mid EFC = efc_0 + \varepsilon] - \lim_{\varepsilon \downarrow 0} E[\mathbf{1}[Pell > 0] \mid EFC = efc_0 + \varepsilon] \right) \\ &+ \tau_1 \left(\lim_{\varepsilon \uparrow 0} E[Pell \mid EFC = efc_0 + \varepsilon] - \lim_{\varepsilon \downarrow 0} E[Pell \mid EFC = efc_0 + \varepsilon] \right) \end{aligned}$$

$$\begin{aligned} &= \tau_0 \left(\lim_{\varepsilon \uparrow 0} E[\mathbf{1}[Pell > 0] \mid EFC = efc_0 + \varepsilon] - \lim_{\varepsilon \downarrow 0} E[\mathbf{1}[Pell > 0] \mid EFC = efc_0 + \varepsilon] \right) \\ &+ \tau_1 \left(\lim_{\varepsilon \uparrow 0} E[Pell \mid EFC = efc_0 + \varepsilon] - \lim_{\varepsilon \downarrow 0} E[Pell \mid EFC = efc_0 + \varepsilon] \right) \end{aligned}$$

Then the RD estimator is equal to:

$$\tau_{RD} = \tau_1 + \tau_0 \left(\frac{\lim_{\varepsilon \uparrow 0} E[\mathbf{1}[Pell > 0] \mid EFC = efc_0 + \varepsilon]}{\lim_{\varepsilon \uparrow 0} E[Pell \mid EFC = efc_0 + \varepsilon] - \lim_{\varepsilon \downarrow 0} E[Pell \mid EFC = efc_0 + \varepsilon]} \right)$$

Where:

$$\frac{\lim_{\varepsilon \uparrow 0} E\left[\mathbf{1}[Pell > 0] \mid EFC = efc_0 + \varepsilon\right]}{\lim_{\varepsilon \uparrow 0} E\left[Pell \mid EFC = efc_0 + \varepsilon\right] - \lim_{\varepsilon \downarrow 0} E\left[Pell \mid EFC = efc_0 + \varepsilon\right]} = \frac{\lim_{\varepsilon \uparrow 0} \Pr\left[\pi = 1 \mid EFC = efc_0 + \varepsilon\right]}{\lim_{\varepsilon \uparrow 0} E\left[\pi \cdot \left(400 - \left(EFC - efc_0\right)\right) \mid EFC = efc_0 + \varepsilon\right]}$$

$$= \frac{\lim_{\varepsilon \uparrow 0} \Pr\left[\pi = 1 \mid EFC = efc_0 + \varepsilon\right]}{400 \lim_{\varepsilon \uparrow 0} \Pr\left[\pi = 1 \mid EFC = efc_0 + \varepsilon\right]} = \frac{1}{400}$$

Thus, as in the sharp case, $\tau_{RD} = \tau_1 + \frac{\tau_0}{Pell(efc_0)}$, where $Pell(efc_0) = 400$. Following the arguments presented in the previous section, and assuming that $f(Pell, \tau)$ does not include any higher order terms, the regression kink estimator is equal to τ_1 and $\tau_0 = (\tau_{RD} - \tau_{RK}) \cdot 400$.

1.10 Figures and Tables

Figure 1.1: Time Series Variation in Maximum Pell Grant Award

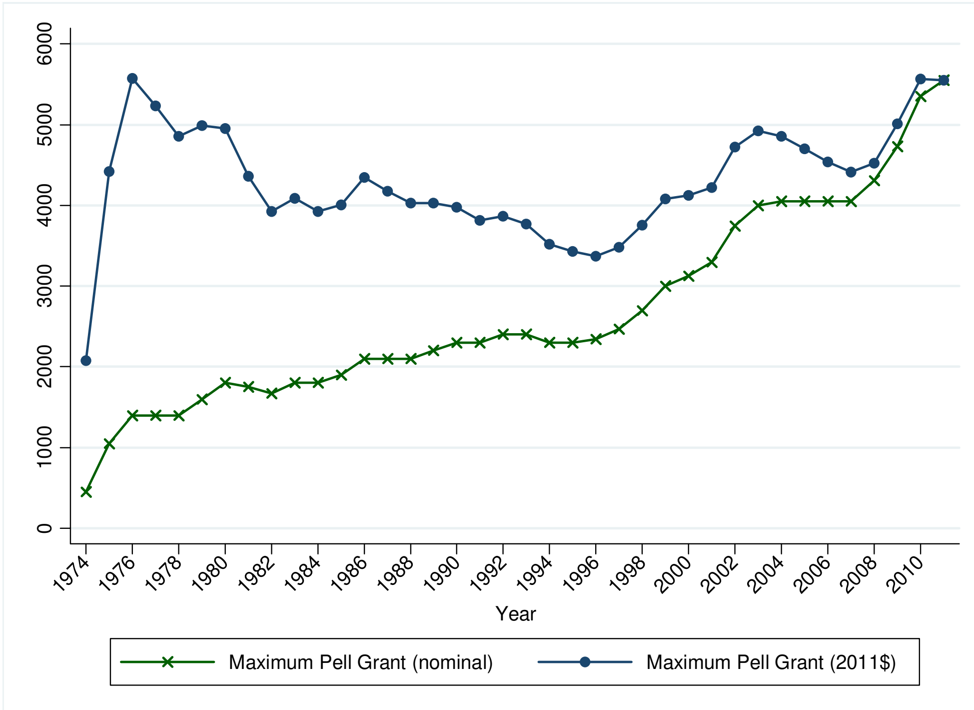


Figure 1.2: The Maximum Pell Grant Award as a Percentage of the Average Cost of Attendance

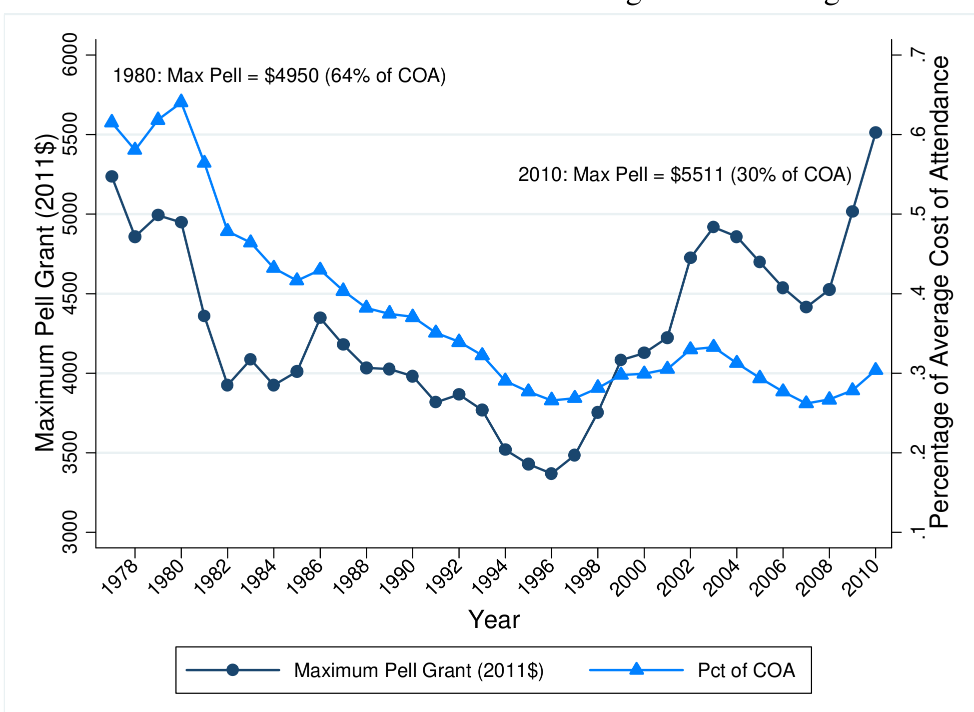
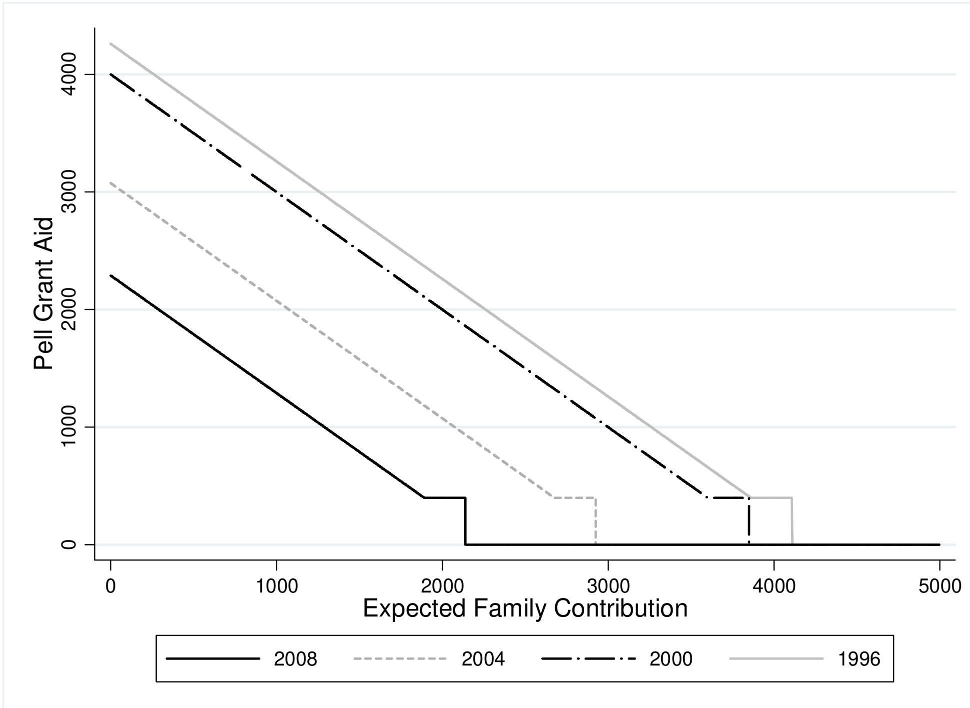
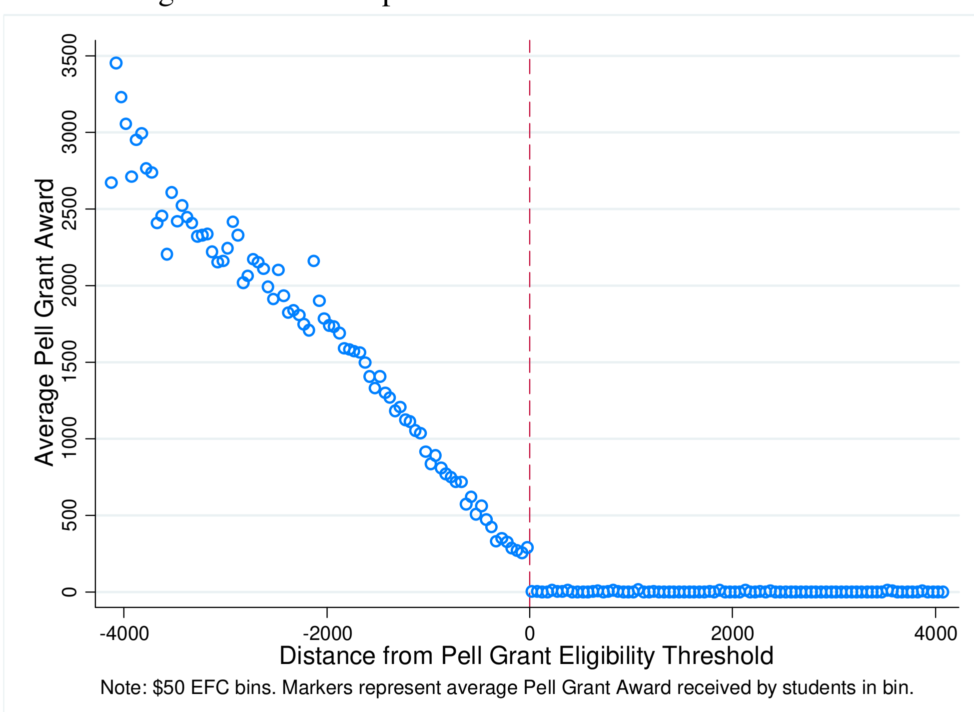


Figure 1.3: Pell Grant Award Schedules, NPSAS Sample Years



Notes: Each line represents the statutory Pell Grant award a full-time, full-year student with a given EFC would receive in the years covered by the NPSAS. All dollar amounts are nominal.

Figure 1.4: The Empirical Distribution of Pell Grant Aid



Notes: $50 EFC bins. Each marker represents the average Pell Grant received by students in the bin.

Figure 1.5: Conceptual Framework, RK/RD Design

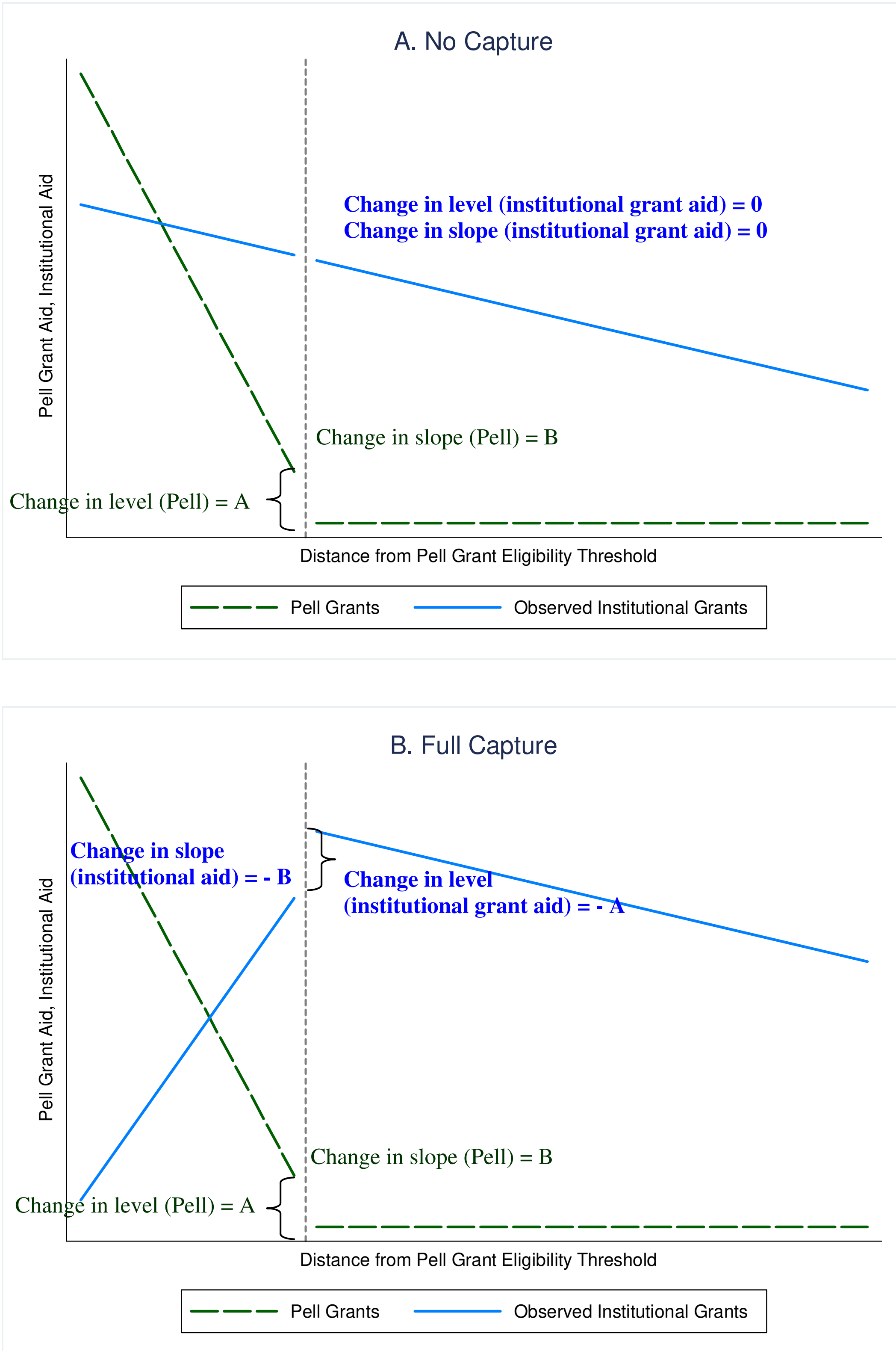
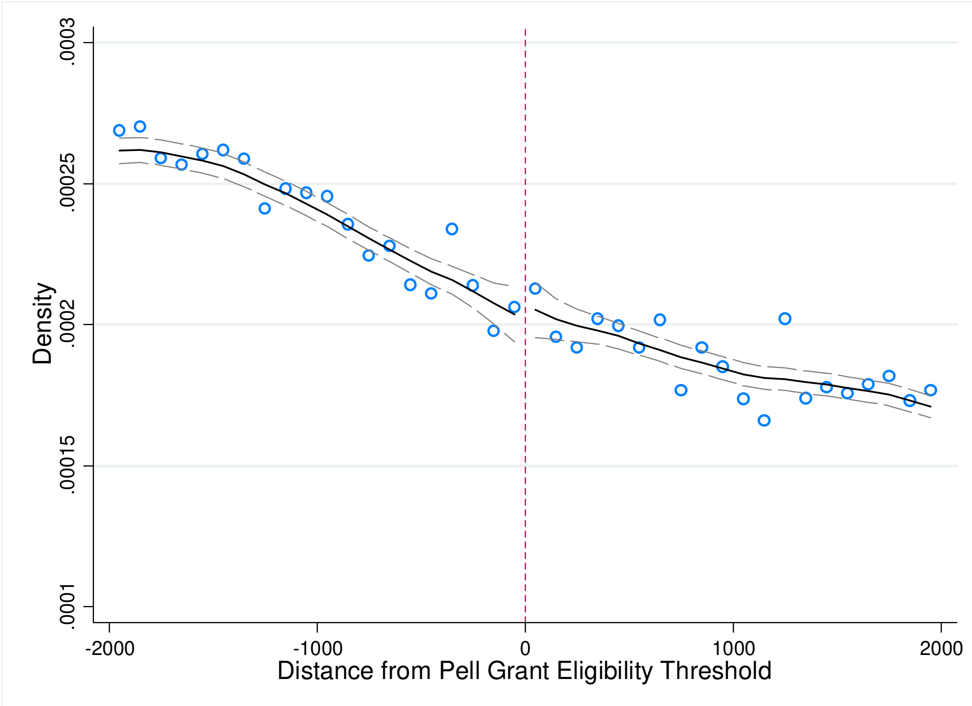
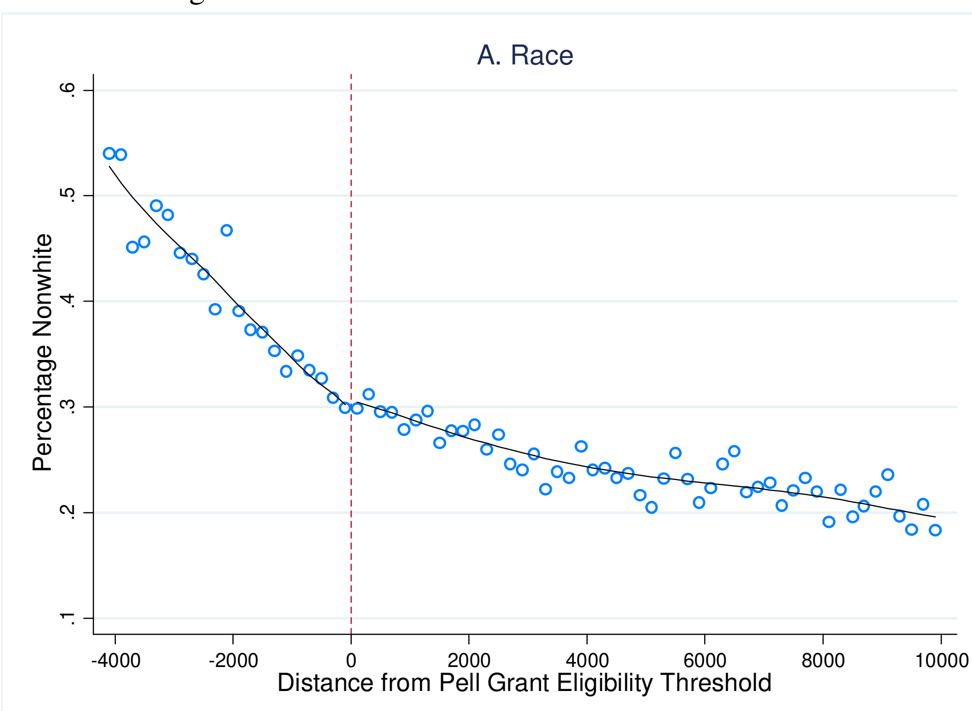


Figure 1.6: The Density of EFC at the Pell Grant Eligibility Threshold



Notes: \$100 EFC bins. Estimated discontinuity (McCrary test) = 0.028 (0.041). Estimated change in slope = -0.113 (0.083).

Figure 1.7: The Distribution of Baseline Covariates



Notes: \$200 EFC bins.

Figure 1.7: The Distribution of Baseline Covariates, cont.

Notes: $200 EFC bins.

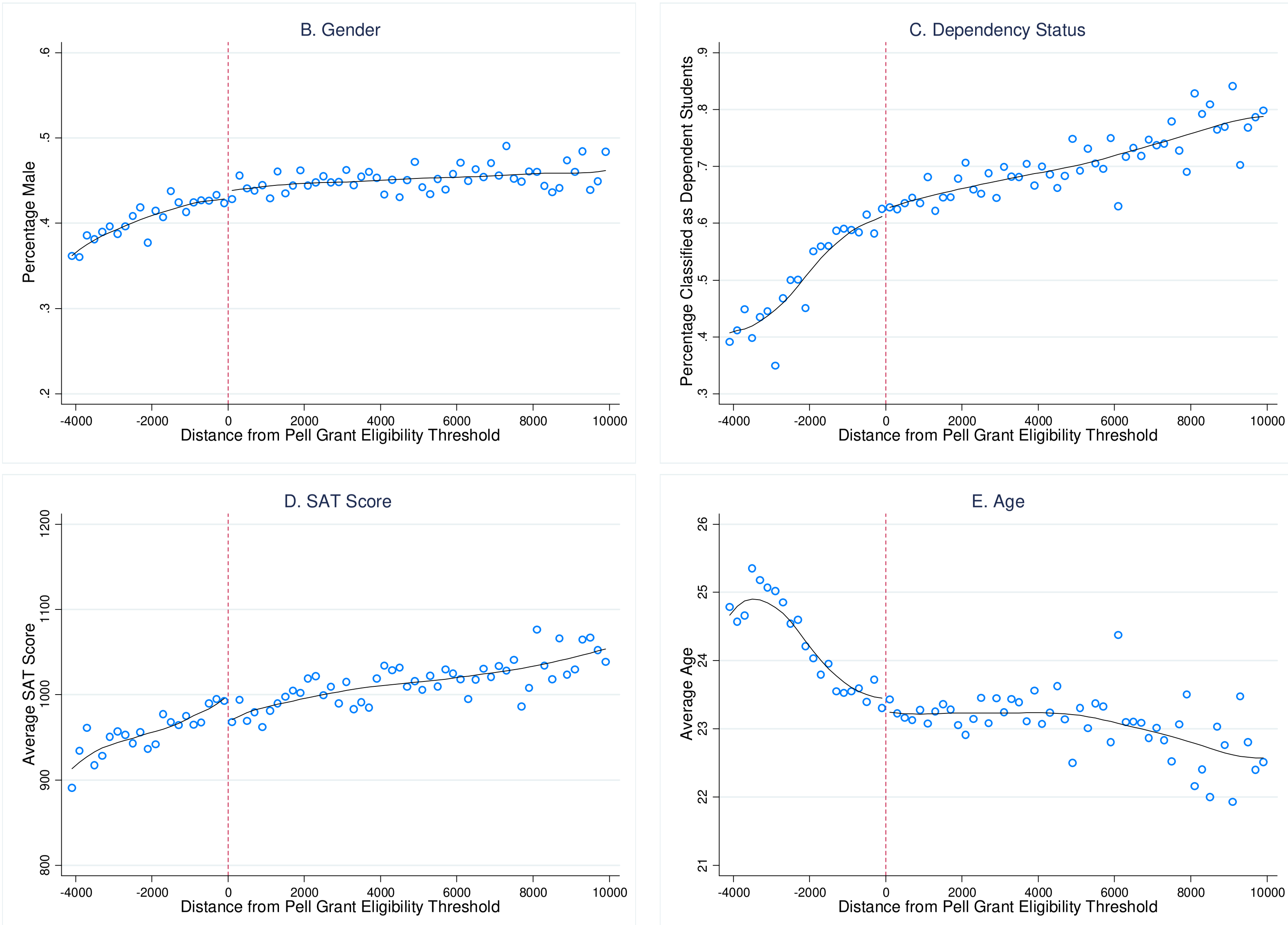


Figure 1.8: The Density of EFC at the Pell Grant Eligibility Cut-off, by Sector

Notes: $100 EFC bins. SAT scores for first-year students only.

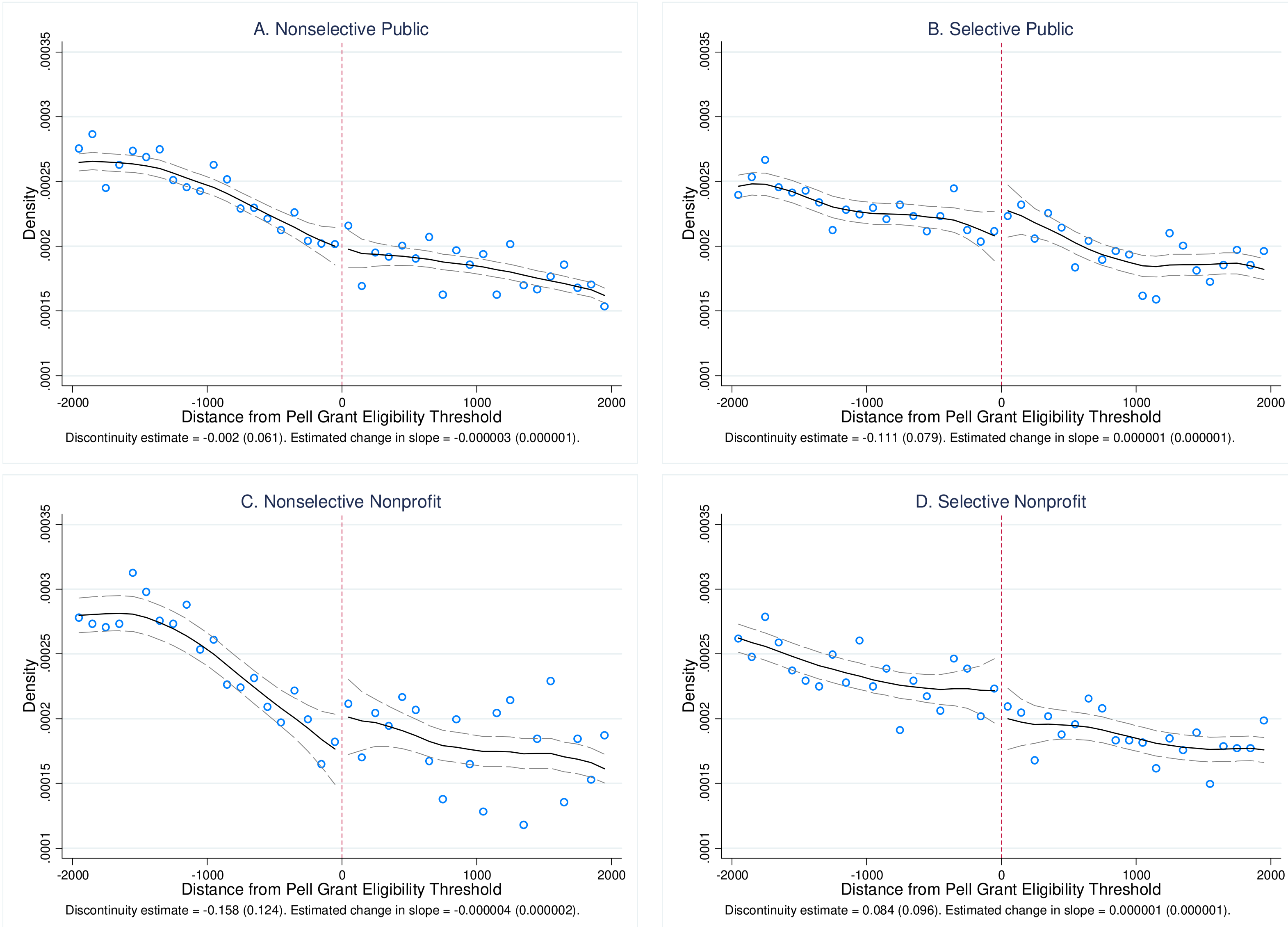
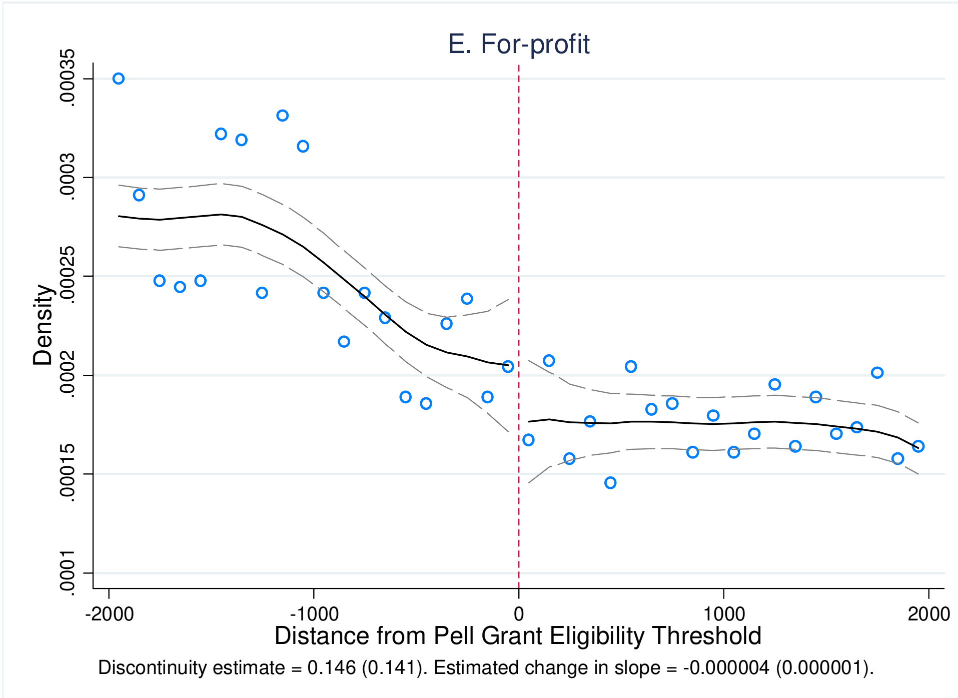
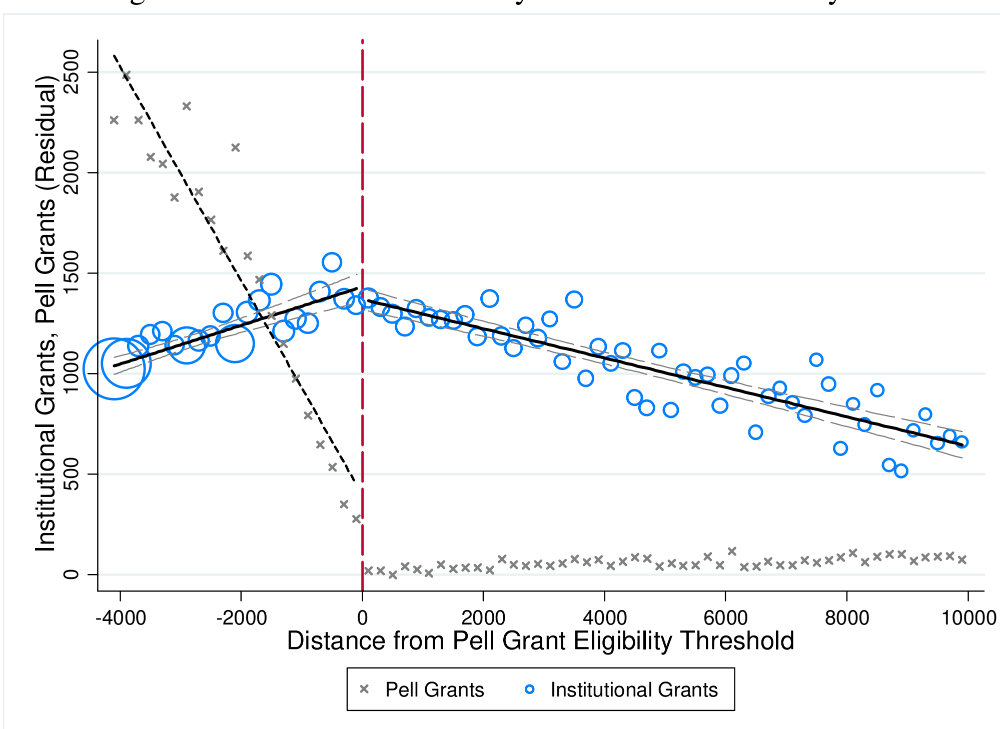


Figure 1.8: The Density of EFC at the Pell Grant Eligibility Cut-off by Sector, continued



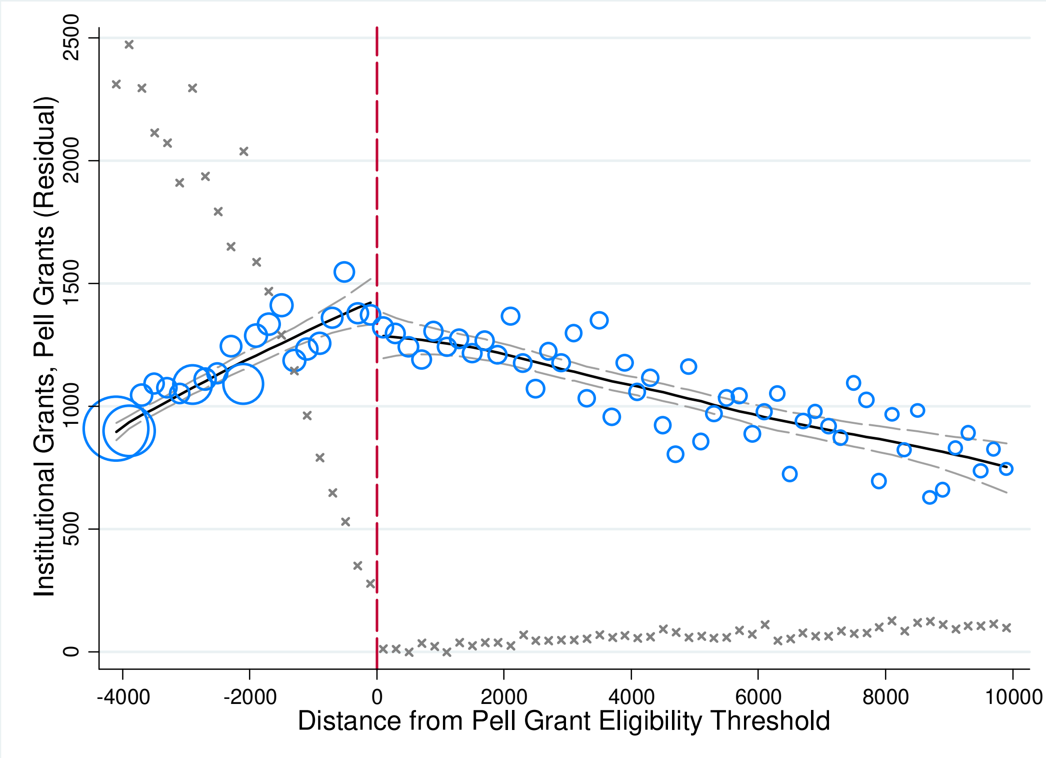
Notes: $100 EFC bins.

Figure 1.9: Pell Grant Generosity and Institutional Aid by EFC



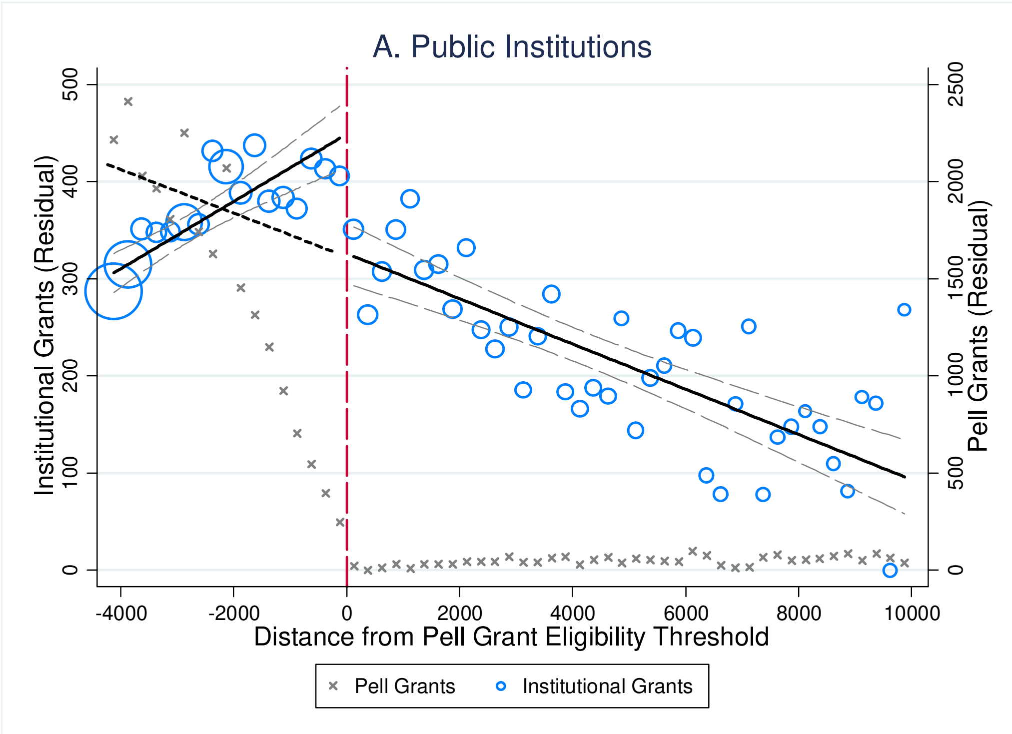
Notes: $200 EFC bins. The black solid line represents a linear fit of institutional grant aid on EFC, estimated separately on each side of the cut-off; gray dashed lines are 95 percent confidence intervals. The thin black dashed line is a linear fit of Pell Grant aid on EFC. Larger circles indicate a larger number of students within the EFC bin.

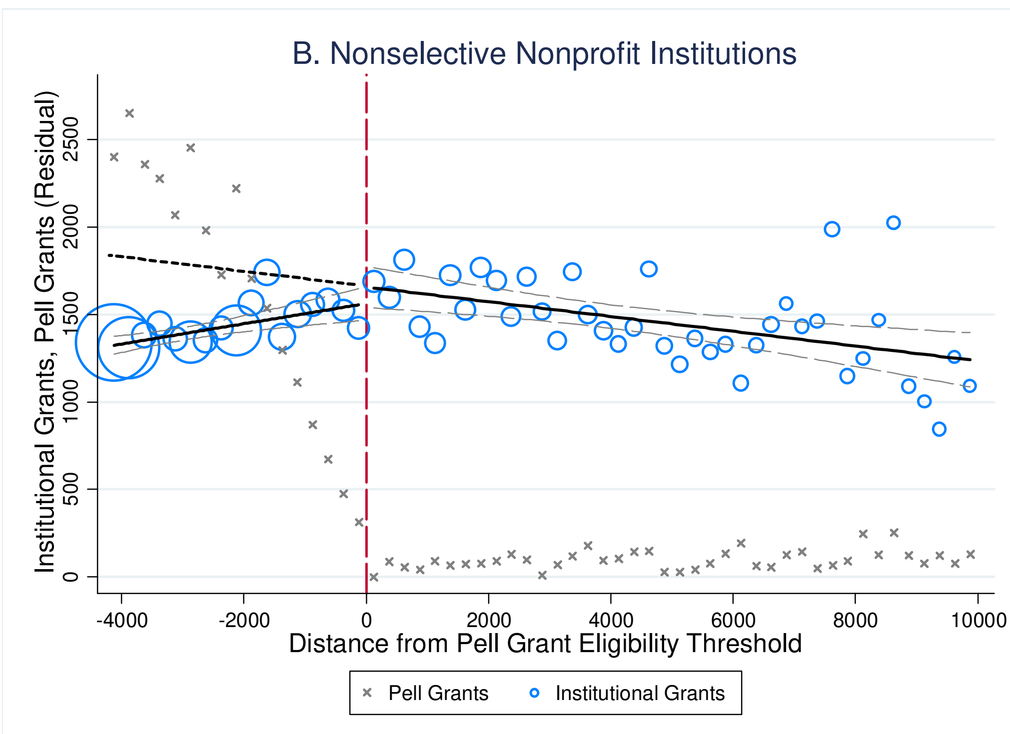
Figure 1.10: Main Results, Local Linear Regression



Notes: $200 EFC bins. The black solid line represents a local linear fit of institutional grant aid on EFC, estimated separately on each side of the cut-off. See Figure 1.9 notes.

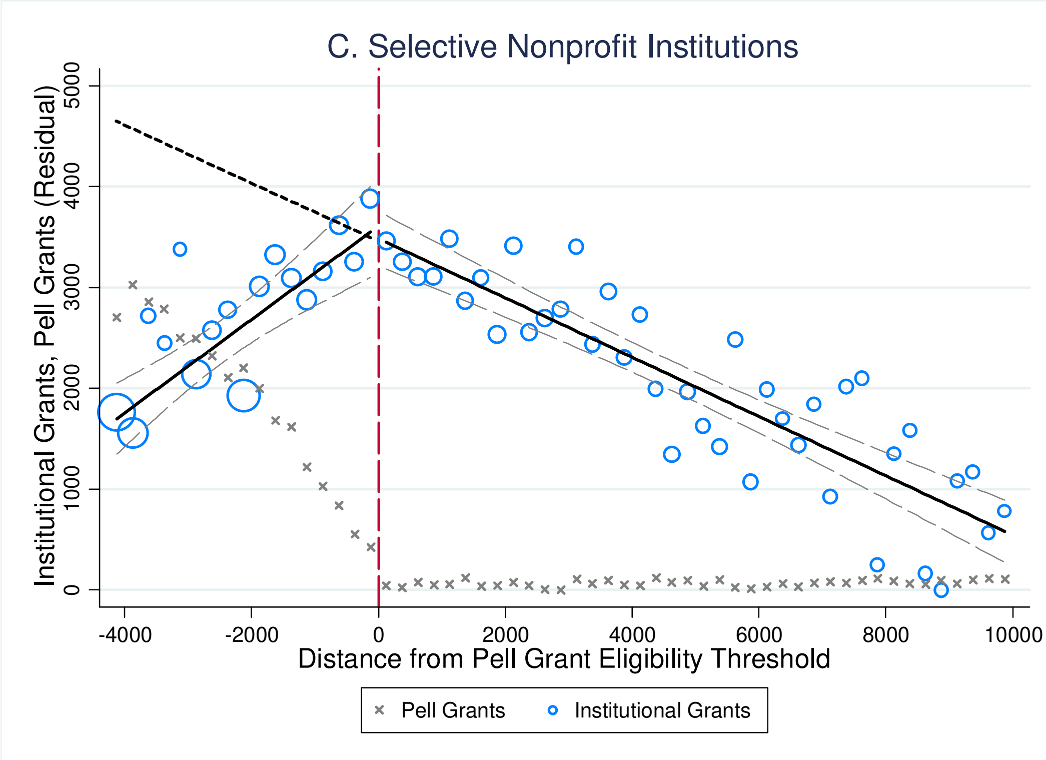
Figure 1.11: Pell Grant Generosity and Institutional Aid by Sector





Notes: $250 EFC bins. The black solid line represents a linear fit of institutional grant aid on EFC, estimated separately on each side of the cut-off; gray dashed lines are 95 percent confidence intervals. The black dashed line is an extension of the linear fit of Pell Grant aid on EFC for Pell ineligible students. Larger circles indicate a larger number of students.

Figure 1.11: Pell Grant Incidence by Sector, continued



Notes: $250 EFC bins. The black solid line represents a linear fit of institutional grant aid on EFC, estimated separately on each side of the cut-off; gray dashed lines are 95 percent confidence intervals. The black dashed line is an extension of the linear fit of Pell Grant aid on EFC for Pell ineligible students. Larger circles indicate a larger number of students.

Figure 1.12: Percentage of Students with any Unmet Need

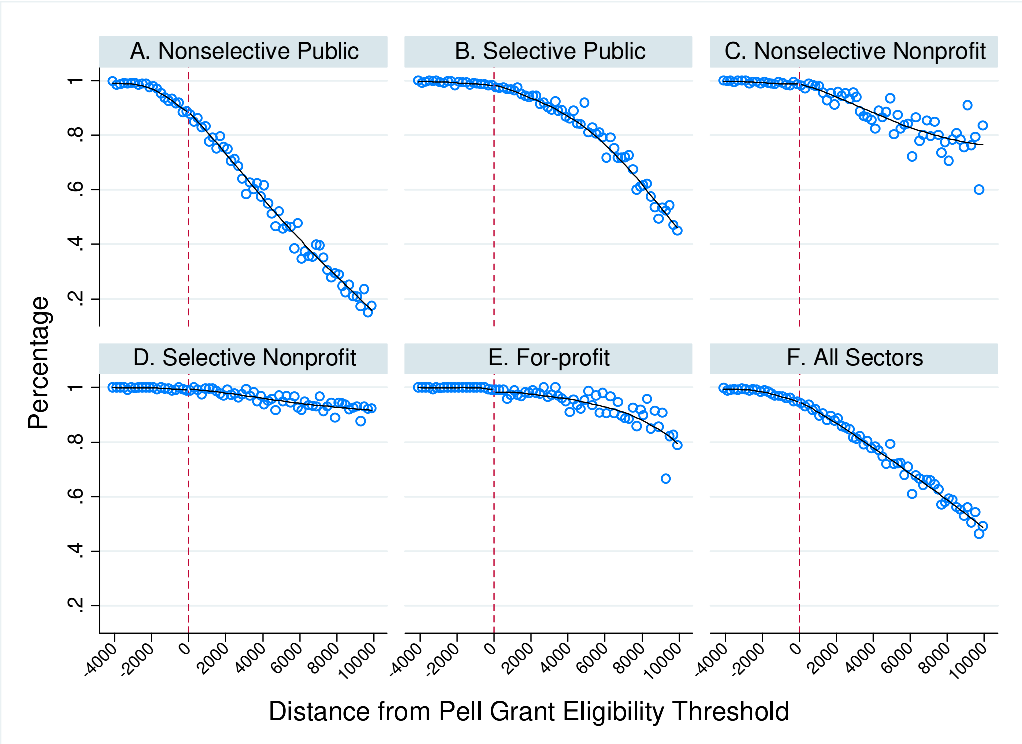
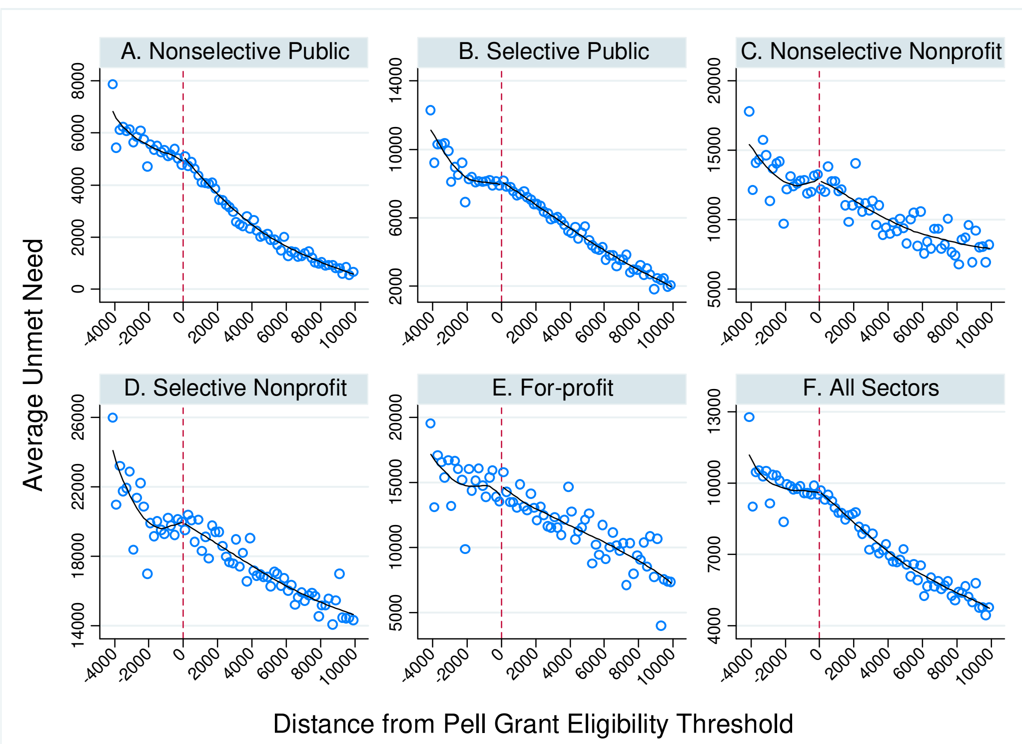
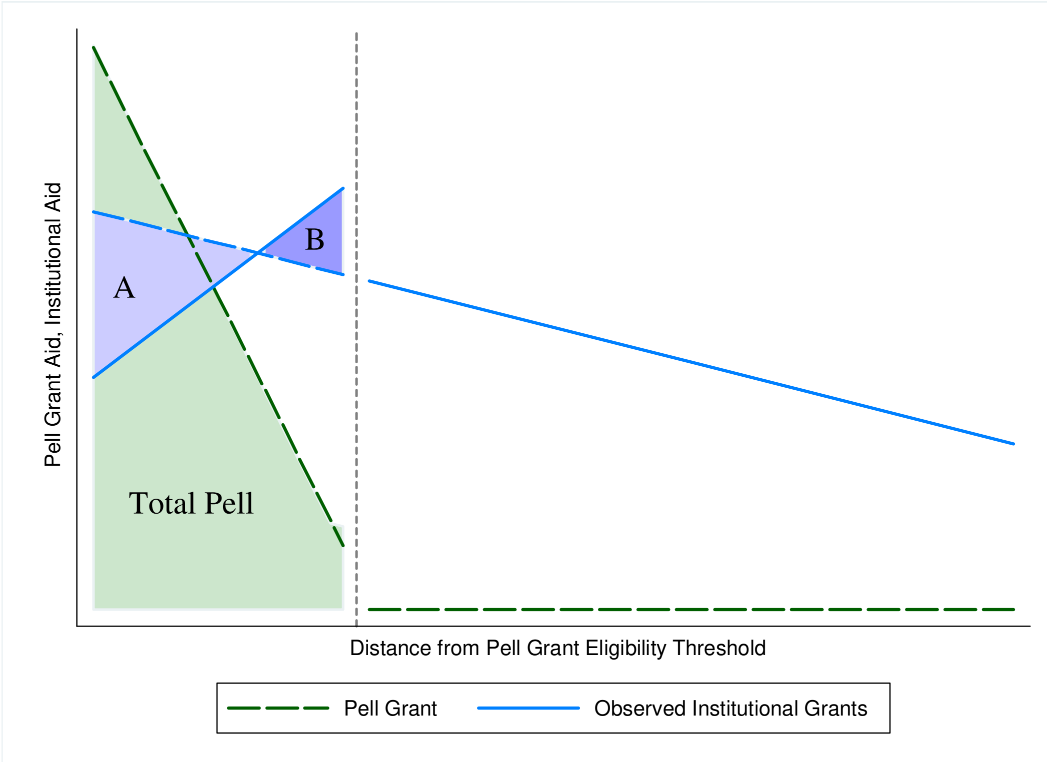


Figure 1.13: Average Unmet Need



Notes: Unmet need equals a student's cost of attendance minus her EFC, Pell Grant and other federal grant aid, and state grant aid.

Figure 1.14: Framework for Estimating the Economic Incidence of the Pell Grant Program



Notes: The area labeled Total Pell represents the total amount of Pell Grant aid disbursed to students. The areas A and B represent the difference between the area below the counterfactual institutional grant aid-EFC relationship (represented by the dashed line) and the actual institutional grant aid-EFC relationship for Pell eligible students (represented by the solid line); A-B represents the amount of institutional grant aid students failed to receive due to the Pell Grant Program. See Section 1.7 for details.

Table 1.1: Characteristics of Schools and Students

	Nonselective Institutions			Selective Institutions		All Schools
	Public	Nonprofit	For-profit	Public	Nonprofit	
Number of Students	55,420	12,350	12,620	34,710	18,160	133,270
Number of Unique Schools	700	260	270	240	340	1,800
Student Financial Aid						
Percentage receiving Pell Grants	0.43	0.50	0.62	0.37	0.38	0.43
Pell Grant aid (nonzero)	$2,768	$2,927	$2,902	$2,899	$2,865	$2,844
Percentage receiving institutional aid	0.13	0.43	0.10	0.26	0.68	0.26
Institutional aid (nonzero)	$1,763	$5,567	$2,865	$3,304	$10,776	$5,949
Net Tuition (tuition - institutional aid)	$2,231	$9,044	$11,607	$4,639	$13,301	$5,887
Student Demographic Characteristics						
Non-white	0.40	0.42	0.52	0.30	0.27	0.37
Male	0.40	0.38	0.42	0.45	0.41	0.41
Dependent student	0.48	0.50	0.32	0.69	0.75	0.56
Age	25	25	26	22	22	24
Expected Family Contribution	$3,368	$3,330	$2,574	$4,267	$4,503	$3,678
Student Attendance Status						
Full-time	0.59	0.79	0.77	0.85	0.89	0.73
Months of enrollment	10.3	10.1	9.7	10.7	10.4	10.3
First-year/freshman	0.48	0.41	0.48	0.25	0.30	0.39

Data: 1996, 2000, 2004, and 2008 NPSAS. **Notes:** Number of observations rounded to nearest 10. All dollar amounts in 2011$. See text for definitions of sectors (public, nonprofit, for-profit, selective, and nonselective). Sample excludes graduate and professional students, students attending multiple institutions during the academic year, students not enrolled in the fall semester, athletic scholarship recipients, noncitizens, and students attending nondegree granting institutions, theological seminaries, or other faith-based institutions.

Table 1.2: Institutional Revenue and Expenditures

	Nonselective Institutions			Selective Institutions		All Schools
	Public	Nonprofit	For-profit	Public	Nonprofit	
Average FTE Students Enrolled	5,590	1,810	1,950	11,390	2,830	4,880
Number of Unique Schools	770	240	290	290	470	2,060
Total Revenue ($100k)	$1,171	$1,042	$3,562	$7,320	$5,681	$3,944
Total Expenditures ($100K)	$1,364	$1,139	$2,308	$9,506	$6,262	$5,681
Revenue - Expenditures ($100K)	-$193	-$97	$1,254	-$2,186	-$581	-$1,737
Pell Grants ($100k)	$87	$44	$488	$113	$29	$106
Pell Grants as a % of Total Revenue	0.07	0.04	0.14	0.02	0.01	0.03

Data: 2000, 2004, 2008 NPSAS and IPEDS. **Notes:** Number of observations rounded to nearest 10; All dollar amounts in 2011$. Sample includes schools serving students described in Table 1 with IPEDS revenue and expenditure data available for 2000, 2004, and 2008.

Table 1.3: Baseline Characteristics, Varying Bandwidths and Polynomials

	1. White		2. Male		3. Dependent		4. SAT Score		5. Age	
	level	derivative	level	derivative	level	derivative	level	derivative	level	derivative
(EFC-k_t) in [-4100,10000]	0.007	0.00004	-0.004	0.00001	0.032	-0.000001	11.57	0.012	-0.260	0.00002
	(0.007)	(0.0001)	(0.006)	(0.00002)	(0.008)**	(0.00001)	(8.17)	(0.009)	(0.104)*	(0.0002)
Optimal Degree of Polynomial	5		1		5		4		5	
Observations	133,270		133,270		133,270		17,080		133,270	
(EFC-k_t) in [-4000,4000]	0.002	0.00004	-0.006	-0.000003	-0.013	-0.0001	9.73	0.043	0.129	0.001
	(0.008)	(0.0001)	(0.010)	(0.00001)	(0.011)	(0.0001)	(9.96)	(0.027)	(0.114)	(0.003)
Optimal Degree of Polynomial	3		3		5		4		5	
Observations	87,310		87310		87310		10240		87310	
(EFC-k_t) in [-3000, 3000]	-0.002	-0.00003	-0.0002	0.00001	-0.009	-0.0001	5.09	0.006	-0.006	0.0002
	(0.009)	(0.00003)	(0.009)	(0.00001)	(0.010)	(0.0002)	(9.09)	(0.005)	(0.102)	(0.0002)
Optimal Degree of Polynomial	4		1		3		1		2	
Observations	62,480		62480		62480		7830		62480	

Data: 1996, 2000, 2004, and 2008 NPSAS. **Notes:** Each cell represents a separate regression. Number of observations rounded to nearest 10. Standard errors clustered at institution level in parentheses; ** p<0.01, * p<0.05, + p<0.1; Pell Grants in constant 2011$. All regressions include year and school fixed effects, and a up to a fifth degree polynominal in student expected family contribution (EFC). Optimal degree of polynomial for each bandwidth determine using the minimum Akaike Information Criterion. SAT scores for first-year students only.

Table 1.4: RK and RD Estimates of the Impact of Pell Grant Generosity on Institutional Aid

	First Stage	Reduced Form	IV (RK)	IV (RD)
	(1)	(2)	(3)	(4)
Change in slope	-0.699	0.153		
	(0.007)**	(0.031)**		
Change in level	397.74	128.45		
	(12.52)**	(42.55)**		
Pell Grant Aid			-0.219	0.323
			(0.044)**	(0.106)**
F-test of excluded instrument(s)			7928	1132
Over-id test (p-value)			0.000	
Observations	133,270	133,270	133,270	133,270

Data: 1996, 2000, 2004, and 2008 NPSAS. **Notes:** Each column represents a separate regression. Number of observations rounded to nearest 10. Standard errors clustered at institution level in parentheses; ** p<0.01, * p<0.05, + p<0.1; Pell Grants and instituional grants in constant 2011$. All regressions include year and school fixed effects, linear and quadratic terms in age, and indicators for gender, race, fall attendance status, enrollment length, level, dependency status, out-of-state student, and a quadratic in student expected family contribution (EFC - k_t, where k_t is the threshold for Pell Grant eligibility in year *t*). In column 3, **1**[EFC < k_t**]** instruments for Pell Grant Aid. In column 4, (EFC - k_t)***1**[EFC<k_t**]** instruments for Pell Grant Aid. Students with EFC greater than 10,000 from Pell Grant eligibility thresold are excluded.

Table 1.5: Robustness of RK and RD Estimates of the Impact of Pell Grant Generosity on Institutional Aid to Varying Bandwidths and Polynomials

	Polynomial of Order:	IV (RK) (1)	IV (RD) (2)
A. (EFC - k_t) in [-4100,10000]	One	-0.294	0.298
		(0.024)**	(0.109)**
	Two	-0.219	0.323
		(0.044)**	(0.106)**
	Three	-0.028	0.315
		(0.070)	(0.174)+
Optimal Degree		2	2
Observations		133,270	133,270
B. (EFC-k_t) in [-4000,4000]	One	-0.173	0.307
		(0.031)**	(0.184)+
	Two	-0.135	0.337
		(0.107)	(0.209)
	Three	-0.153	0.438
		(0.110)	(0.475)
Optimal Degree		1	1
Observations		87,290	87,290
C. (EFC-k_t) in [-3000, 3000]	One	-0.183	0.383
		(0.047)**	(0.289)
	Two	-0.188	0.435
		(0.134)	(0.323)
	Three	-0.208	0.973
		(0.142)	(1.147)
Optimal Degree		1	1
Observations		62,420	62,420

Data: 1996, 2000, 2004, and 2008 NPSAS. **Notes:** Each cell represents a separate regression. Number of observations rounded to nearest 10. Standard errors clustered at institution level in parentheses; ** p<0.01, * p<0.05, + p<0.1; Pell Grants and instituional grants in constant 2011$. All regressions include year and school fixed effects, linear and quadratic terms in age, and indicators for gender, race, fall attendance status, enrollment length, level, dependency status, out-of-state student, and a up to a third degree polynominal in student expected family contribution (EFC - k_t, where k_t is the threshold for Pell Grant eligibility in year t). The optimal degree of polynomial for each bandwidth is determined using the minimum Akaike Information Criteria. RD estimates instrument for Pell Grant aid with **1**[EFC < k_t]**;** RK estimates instrument with (EFC - k_t)***1**[EFC < k_t]**.**

Table 1.6: The Impact of Pell Grant Generosity on Institutional Aid, Treatment Dimensions

	Pass-Through	Willingness to Pay
A. All institutions	-0.219	260.5
	(0.044)**	(50.06)**
Observations	133,270	
B. By sector		
Public Nonselective	-0.179	318.3
	(0.017)**	(63.31)**
Public Selective	-0.173	618.9
	(0.032)**	(101.5)**
Nonprofit Nonselective	-0.154	-193.3
	(0.060)*	(216.6)
Nonprofit Selective	-0.687	97.23
	(0.101)**	(248.3)
For-profit	-0.133	84.67
	(0.029)**	(80.84)
Observations	133,270	

Data: 1996, 2000, 2004, and 2008 NPSAS. **Notes:** Each column within a panel represents a separate regression. Number of observations rounded to nearest 10. Standard errors clustered at institution level in parentheses; ** p<0.01, * p<0.05, + p<0.1. Pell Grants and instituional grants in constant 2011$. All regressions include year and school fixed effects, linear and quadratic terms in age, and indicators for gender, race, fall attendance intensity, enrollment length, level, dependency status, out-of-state student, and a linear term in student expected family contribution. Panel A also includes a quadratic in EFC. Students with EFC greater than 10,000 from Pell Grant cut-off are excluded. See text for definitions of treatment dimensions.

Table 1.7: Heterogeneity in the Impact of Pell Grant Generosity on Institutional Aid by Sector and Demographics

	Nonwhite	White	Independent	Dependent	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Public						
Pass-through	-0.207	-0.183	-0.073	-0.232	-0.208	-0.195
	(0.031)**	(0.021)**	(0.013)**	(0.024)**	(0.021)**	(0.027)**
Willingness to pay	670.6	338.6	361.8	471.5	452.7	465.2
	(115.9)**	(50.19)**	(88.61)**	(74.89)**	(61.01)**	(79.16)**
Private Nonselective						
Pass-through	-0.134	-0.150	-0.009	-0.171	-0.145	-0.165
	(0.047)**	(0.0500)**	(0.030)	(0.053)**	(0.044)**	(0.049)**
Willingness to pay	-27.17	-68.76	-147.3	-185.3	-11.58	-136.1
	(159.0)	(142.0)	(116.5)	(196.2)	(139.6)	(164.6)
Nonprofit Selective						
Pass-through	-0.438	-0.982	0.144	-0.609	-0.665	-0.716
	(0.163)**	(0.138)**	(0.128)	(0.115)**	(0.131)**	(0.146)**
Willingness to pay	-704.5	441.5	-505.2	18.36	-117.2	373.9
	(704.6)	(256.6)+	(375.6)	(309.7)	(339.6)	(367.3)
Observations	49,360	83,910	59,090	74,180	78,140	55,130

Data: 1996, 2000, 2004, and 2008 NPSAS. **Notes:** Each column represents a separate regression. Number of observations rounded to nearest 10. Standard errors clustered at institution level in parentheses; ** p<0.01, * p<0.05, + p<0.1. Pell Grants and instituional grants in constant 2011$. All regressions include year and school fixed effects, linear and quadratic terms in age, and indicators for gender, race, fall attendance status, enrollment length, level, dependency status, out-of-state student, and a linear term in student expected family contribution. Students with EFC greater than 10,000 from Pell Grant cut-off are excluded. See text for definitions of treatment dimensions.

Table 1.8: Heterogeneity in Pass-Through by Market Concentration

	(1) Baseline	(2) All Competitors		(3) Direct Competitors	
		Unconcentrated	Concentrated	Unconcentrated	Concentrated
A. All institutions	-0.189	-0.227	-0.169	-0.187	-0.188
	(0.047)**	(0.057)**	(0.046)**	(0.065)**	(0.046)**
Test of equality (p-value)		*0.129*		*0.987*	
Observations	108,400	108,400		108,400	
B. By sector					
Public	-0.106	-0.150	-0.087	-0.157	-0.094
	(0.048)*	(0.052)**	(0.048)+	(0.057)**	(0.047)*
Nonselective Private	-0.112	-0.109	-0.116	-0.117	-0.109
	(0.053)*	(0.058)+	(0.068)+	(0.060)+	(0.064)+
Selective Nonprofit	-0.719	-0.699	-0.747	-0.442	-0.794
	(0.112)**	(0.175)**	(0.130)**	(0.247)+	(0.118)**
Test of equality (p-value):					
Public		*0.011*		*0.039*	
Nonselective Private		*0.918*		*0.901*	
Selective Nonprofit		*0.812*		*0.169*	
Observations	108,400	108,400		108,400	

Data: 2004 and 2008 NPSAS, 2003 and 2007 IPEDS. **Notes:** Each column within a panel represents a separate regression. Number of observations rounded to nearest 10. Standard errors clustered at institution level in parentheses; ** p<0.01, * p<0.05, + p<0.1. Pell Grants and instituional grants in constant 2011$. All regressions include year and school fixed effects, linear and quadratic terms in age, and indicators for gender, race, fall attendance status, enrollment length, level, dependency status, out-of-state student, and a linear term in student expected family contribution (EFC). Panel A also includes a quadratic in EFC. Students with EFC greater than 10,000 from Pell Grant cut-off are excluded. A school's market is considered concentrated if the Herfindahl index of institutional FTE undergraduate student shares is greater than 0.25. In column 2, schools in all sectors in a given market are considered competitors. In column 3, only schools with similar selectivity in a market are considered competitors.

Table 1.9: RK Estimates of the Impact of Pell Grant Aid on Institution Quality

	Tuition/FTE	Revenue/FTE	Institutional Expenditures/FTE on: Grants	Instruction	Student Services	CDR
	(1)	(2)	(3)	(4)	(5)	(6)
A. All Sectors						
Mean of depvar	$10,619	$19,038	$1,061	$6,214	$5,748	6.55
* Pell Grant Aid	-0.027	-0.030	0.004	-0.035	0.008	0.0003
	(0.142)	(0.198)	(0.015)	(0.060)	(0.072)	(0.0001)**
Observations	66,950	77,470	66,940	83,810	84,630	128,800
B. By Sector						
Nonselective Public						
Mean of depvar	$5,160	$13,629	$1,086	$5,051	$3,828	8.2
* Pell Grant Aid	-0.089	-0.153	0.014	-0.023	-0.047	-0.0001
	(0.040)*	(0.074)*	(0.008)+	(0.024)	(0.025)+	(0.00007)
Selective Public						
Mean of depvar	$7,839	$25,364	$1,495	$8,412	$5,503	4.5
* Pell Grant Aid	0.082	0.070	0.022	0.139	0.034	0.0003
	(0.059)	(0.170)	(0.013)+	(0.064)*	(0.036)	(0.0001)**
Nonselective Nonprofit						
Mean of depvar	$15,247	$22,260	$799	$6,138	$7,872	7.1
* Pell Grant Aid	-0.043	0.120	0.033	0.008	0.116	0.0001
	(0.155)	(0.259)	(0.030)	(0.088)	(0.093)	(0.0004)
Selective Nonprofit						
Mean of depvar	$22,449	$32,393	$1,500	$9,489	$10,288	3.2
* Pell Grant Aid	0.088	0.071	0.038	-0.038	0.064	0.0003
	(0.175)	(0.288)	(0.051)	(0.097)	(0.112)	(0.0001)**
For Profit						
Mean of depvar	$14,409	$15,860	$353	$3,522	$8,545	10.1
* Pell Grant Aid	-0.231	-0.277	-0.006	0.022	-0.228	-0.001
	(0.133)+	(0.161)+	(0.013)	(0.057)	(0.156)	(0.0003)**
Observations	66,950	77,470	66,940	83,810	84,630	128,800

Data: 1996, 2000, 2004, and 2008 NPSAS, 2003 and 2007 IPEDS, Department of Education Official Cohort Default Rates. **Notes:** Each column within a panel represents a separate regression. Standard errors clustered at institution level in parentheses; ** $p<0.01$, * $p<0.05$, + $p<0.1$. Columns 1 through 6 include students attending institutions in 2004 and 2008 with revenue or expenditure information available in prior year IPEDS. Column 6 includes students attending institutions in all years with information on two-year cohort default rates. Number of observations rounded to nearest 10. Regressions include year fixed effects and a linear term in student expected family contribution (EFC). Panel A also includes a quadratic in EFC.

Table 1.10: The Incidence of Pell Grant Aid across all Recipients

	Percent Captured	95% CI
All Institutions	0.163	[0.114, 0.212]
Public Institutions	0.031	[0.002, 0.060]
Nonselective Private Institutions	0.176	[0.062, 0.290]
Selective Nonprofit Institutions	0.787	[0.563, 1.011]

Data: 1996, 2000, 2004, and 2008 NPSAS. **Notes:** These estimates assume the observed institutional aid-EFC relationship for Pell ineligible students is a valid counterfactual for Pell eligible students in the absence of the Pell Grant Program. The overall percentage of Pell Grant aid captured by institutions is equal to the ratio of the difference between the area below the counterfactual Pell Grant-EFC curve and the actual Pell Grant-EFC curve and the overall transfer of Pell Grant aid to eligible students (refer to section 6 for details).

CHAPTER 2

The Returns to Higher Education for Marginal Students

2.1 Introduction

Public two-year colleges serve an increasingly important role in meeting the growing demand for an educated workforce in the United States. The two-year sector of higher education absorbed much of the growth in college attendance over the past several decades and students induced to go to college by changes in the costs or returns to higher education are more likely to attend a two-year institution (Bound et al. 2010). Numerous studies provide evidence of substantial labor market returns to community college credits and credentials, with estimated earnings gains as high as 13 percent for each year of attendance and 30 percent for associate degree receipt (e.g., Kane and Rouse 1995, Jacobson et al. 2005). However, the marginal individual induced to enter college by changes in costs or returns may differ substantially from the average college student and it is unclear whether these individuals will experience equally large returns from attending a two-year institution.

In this paper, I focus on a group of students who are likely especially constrained in their ability to finance college attendance – mothers who are current and former welfare recipients in Colorado. Mothers at risk for welfare receipt are especially relevant group given their generally low levels of income, education, and limits on lifetime cash assistance. Using information on earnings trajectories before and following entry into Colorado's welfare program and taking advantage of variation in pre-existing county policies that affect the cost of college-going for several cohorts of women, I estimate medium term impacts of community college attendance, credits, and credential receipt on employment and earnings.

I find that women who are induced to attend college following welfare entry experience large and significant earnings gains, however, these effects are primarily driven by credentials receipt. Women appear to benefit from all community college credentials, including short-term certificates and career-oriented associate's degrees; the sole exception to this finding associate of arts or general studies degrees. These credentials, while potentially facilitating transfers to a four-year program, do not appear to lead to increases in employment or earnings alone.

My identification strategy uses an event-study framework to deal with concerns of selection bias. If college-going women only differ due to unobservable characteristics, such as ability or motivation, these estimates will represent the causal impact of higher education on labor market outcomes. Consistent with prior studies, my results suggest that women earning degrees in health, science, or technical fields experience the largest benefits on the labor market, but even those who earn short-term certificates in non-technical fields experience an increase in earnings following degree receipt. I document that failure to account for sub-associate credentials results in falsely attributing positive earnings gains to college attendance in the absence of degree receipt. Measures of educational attainment in most major surveys (e.g., decennial census, current population survey) do not include certificate receipt, suggesting the large category generally classified as "some college" includes a heterogeneous group of individuals.[1]

Finally, I use information on direct and indirect costs of college attendance and reliance on public assistance to illustrate the potential short-run private and social returns college attendance. In the short-run, when foregone earnings and the direct costs of college attendance are taken into consideration, the private rate of return to certificates and most degrees is negative, providing a rationale as to why so few women complete credentials, even in light of the large impacts on earnings in the medium-term. I also find suggestive evidence of small negative (albeit marginally significant) impacts on welfare receipt in the short-run.

[1] For instance, in the 2000 Census, 22 percent of adults belonged to this category.

My findings speak to the question of whether state and federal welfare policy should support formal human capital development. Both the Obama administration and the Gates Foundation have directed substantial attention and funds towards community college.[2] Additionally, community college students are more likely to benefit from increases in federal aid generosity. Grant aid provided through the Pell Grant Program, the largest source of need-based aid in the United States, grew from $7.2 to $30 billion between 2000 and 2010, with the percentage of recipients attending community colleges increasing from 36 to 57 percent.[3] My results suggest that supports for credential completion are as, if not more important than funding directed towards increasing college attendance.

The remainder of this chapter proceeds as follows: in Section 2.2, I discuss the evolution of federal welfare policy towards human capital development, while the third section focuses on research findings on the returns to the two-year sector of higher education. Section 2.4 discusses my data and presents descriptive results, Section 2.5 discusses my primary empirical strategy and resulting estimates, and the sixth section concludes.

2.2 Welfare Policy and Human Capital Development

Over the past two decades, federal welfare policy has become increasingly less supportive of college enrollment for parents receiving cash assistance. Prior to the large federal reforms enacted in 1996, welfare recipients' behavior was relatively unconstrained under the Aid to Families with Dependent Children (AFDC) program. Attending a post-secondary institution

[2] The 2010 Health Care and Education Reconciliation Act directed $2 billion in competitive grants to community colleges while the 2009 American Recovery and Reinvestment Act included substantial funding for Pell Grants, workforce training programs, and work-study funds in community colleges. In October 2010, the Gates Foundation launched the Completion by Design program, which will award $35 million in competitive grants to community colleges in nine states.

[3] 1999-2000 and 2009-2010 Federal Pell Grant Program End-of-Year Reports, U.S. Department of Education, Office of Postsecondary Education

while on welfare was not uncommon prior to welfare reform – one study estimates that 14 percent of all recipients attended college while on welfare prior to 1996 (London 2006).

The 1996 Personal Responsibility and Work Reconciliation Act (PRWORA) made it more difficult to receive welfare concurrently with college attendance; the legislation imposed work requirements and time limits on welfare recipients, increasing the costs associated with post-secondary education by requiring part-time employment or exit from welfare for those wishing to attend college full-time. Although states had the option to exempt welfare recipients enrolled in college from work requirements in their first two years of receipt, the policy change led to a significant reduction in college enrollment for single mothers. Dave et al. (2008) estimate that PRWORA decreased the probability of college enrollment by 20 percent among all adult mothers. An additional study finds similar impacts on full-time vocational education (Dave et al. 2011).

The 2005 Deficit Reduction Act (DRA), increased pressure on states to place welfare recipients in “employment-related activities”. PROWRA required states to have at least 50 percent of single parents participating in a specific set of work activities, including paid employment, workfare/community service, and education/training. Before DRA, states could deduct each percentage point of caseload reduction since 1995 from their work participation rate target. Since caseloads declined steeply following PRWORA, for most states, including Colorado, this requirement was not binding. DRA reset the base year to 2005. Furthermore, DRA Recipients were further limited to a total of 12 lifetime months of education that must be combined with another, employment-related activity, essentially limiting individuals to part-time college attendance while receiving welfare.

DRA reduced concurrent college attendance and cash assistance receipt for women in Colorado by over 20 percent. Figure 2.1 plots the percentage of recipients also attending a public college in each month prior to and following the policy change, and shows while 9 to 11 percent of women were enrolled in college while receiving assistance before the policy change, only around 7.5 percent were attending college after the rules were implemented. Conversely, total female enrollment in Colorado's community colleges increased 9 percent between the fall of 2004 and fall of 2009.[4]

Whether low-skill parents at risk for welfare receipt benefit from higher education is an open question. The majority of studies that explicitly focus on welfare recipients compare the returns to education vs. a "work first" approach (e.g., job search assistance) in terms of income and welfare receipt. Earlier evaluations found returns to education and human capital development activities to be small or non-existent (e.g., findings from California's GAIN experiment (Riccio and Friedlander 1992; Riccio et al. 1994)); later research illustrated the possibility that any earnings increases following education might be slower to appear. For instance, Hotz et al. (2006) find that an education-focused approach led to greater long-term earnings gains compared to a work-first approach among California welfare recipients. Although these studies offer the advantage of randomization, the "treatment" generally includes types of education and training activities, including basic skills training, vocational training, and community college coursework, with limited information on the amount of time spent in these activities and whether training or courses were attended and what, if any, degrees were received.

[4] Colorado Department of Higher Education, "Student Headcount by Gender: Colorado Public Two-Year Institutions of Higher Education" reports.

2.3 Prior Research

Numerous papers estimate positive returns to both college attendance and credential receipt in the two-year sector. Kane and Rouse (1995) find that each year of community college attendance is correlated with a 7 percent increase in annual earnings and women who complete two-year associate's degrees earn 30 percent more than high school graduates. However, results from cross-sectional studies that compare labor market outcomes across individuals with different levels of education at a point in time may be biased by unobservable characteristics that are correlated with both the decision to go to college and employment or earnings (Card, 1999). Additionally, the data used in these studies does not measure certificate receipt or differentiate between academic and nonacademic associate's degrees.

One approach used to dealing with this type of selection bias is to use repeated observations of earnings and employment for individuals who are employed both before and after they enroll in college. With a sufficiently long panel of data, this event study framework essentially uses a student's pre-college earnings as a counterfactual for what she would have earned had she not gone to school, eliminating concerns of selection on time-invariant characteristics such as ability or motivation. This strategy necessarily excludes "traditional" students who matriculate immediately following high school graduation; however, community college students increasingly older, non-traditional students. Additional federal workforce retraining programs focus their services on older, displaced workers, and among the women in my sample who attend college following welfare entry, the average age is 28.

Jacobson et al. (2005) first use this approach to examine the returns to community college enrollment for high tenure, displaced workers. They estimate that each year of college attendance leads to a 13 percent increase in women's earnings. The benefits of community college attendance primarily accrue to individuals who take technical coursework. Their data only measures degree receipt and not receipt of credentials, such as short-term certificates, and few individuals in their sample remain in college long enough to complete an associate's degree.

Along with my paper, Jepsen et al. (2009) produce some of the first estimates of the returns to sub-associate credentials that do not rely on cross-sectional variation. Their paper focuses on two cohorts of degree-seeking community college students in Kentucky and also uses an individual fixed effects approach to deal with selection on unobservables. Because their sample only includes college-going individuals, returns are measured relative to the average credits earned by drop-outs, rather than relative to no college attendance. Women earnings increase by over 50 percent after earning an associate's degree. Certificates requiring two to three semesters of coursework result in similar earnings gains, while the returns to short-term certificates are small (4 percent).[5]

2.4 Data and Descriptive Results

My underlying sample consists of adult welfare recipients who entered Colorado's welfare system, Colorado Works, between the third quarter of 2004 and second quarter of 2007. I use program data which contains the universe of individuals who received any assistance from Colorado's TANF program (Colorado Works) beginning September 2004. I observe monthly welfare receipt between entry and the first quarter of 2010 as well as a variety of characteristics including age, race, marital status, number of children, lifetime months of assistance, whether the individual is listed as having a disability, and whether she owns a vehicle. Moreover, I observe

[5] Cellini and Chaudhary (2011) also use an event-study approach to test whether individuals attending private two-year institutions experience greater returns than those attending public institutions. The authors find similar returns to public and private schools – around a 6 percent increase in earnings for each year of attendance and a 15 to 17 percent increase in earnings for those that earn an associate's degree.

changes in time-varying individual characteristics (e.g., marriages, births, becoming disabled) that are likely correlated with both the decision to enroll in college and labor market outcomes.[6]

Individuals are linked to information from the Colorado Department of Higher Education which covers individuals' lifetime college attendance, program of study, cost of attendance, financial aid, and credit and degree receipt at all public two and four-year colleges in Colorado. I convert semesters attended to quarters, with the fall semester corresponding with the fourth quarter, winter/spring semester corresponding to first and second quarters, and the summer semester corresponding to the third quarter of a year. Individuals are also matched to data on quarterly earnings and employment in all covered sectors from the Colorado Department of Labor for the 29 quarters between the third quarter of 2003 and the fourth quarter of 2010. Finally, I merge quarterly county unemployment rates from the Bureau of Labor Statistics' Local Area Unemployment Statistics program to the panel.

I drop the small number of male recipients and women who were younger than 19 or older than 60 at entry. My sample includes the small number women who attended four-year institutions, either prior to or following welfare receipt, although my results are robust to excluding these individuals. The remaining sample includes unique 29,556 individuals or 857,124 person-quarter observations.

2.4.1 Characteristics of Recipients by Post-entry College Attendance

Table 2.1 displays the characteristics of the women in my sample, distinguishing between women who do not enroll in college following welfare entry, women who enroll but did not obtain a credential, and credential recipients. The fourth column includes women who were already enrolled in college for at least two quarters at entry rather than those who were induced

[6] Unfortunately, I only observe these time-varying characteristics when individuals are receiving welfare benefits. For individuals who leave welfare, I use the last reported value of these characteristics. If individuals leave and return to welfare, I smooth the imputation of characteristics between spells.

to enter college after entering Colorado's welfare program (however, their pre-welfare labor market attachment is no weaker than that of women in the other groups). College-going women are slightly younger, are more likely to own vehicles, and are slightly less likely to be disabled. A portion of women in all four groups have attended college in the past. For instance, 34 percent of women who attend college and earn a degree following welfare entry enrolled in college for at least one semester. Finally, women who are induced to attend college following entry have higher pre-welfare earnings and employment, suggesting that estimates that do not account for individual effects may be biased upwards.

The bottom portion of Table 2.1 contains information pertaining to the educational attainment and labor market outcomes I examine. Degree recipients earn close to 70 credits after welfare entry. The majority of these women earn a short-term vocational certificate, generally requiring 15 to 30 credits, or one to two semesters of part-time attendance, 14 percent earn a certificate requiring two to three semesters of full-time attendance, and 28 percent earn an associate's degree.[7] Associates' degrees require between 60 and 90 credits of coursework. Associate of arts and associate of general studies (AA/AGS) degrees are designed for students who intend to transfer to a four-year program, and while these degrees may be awarded in specific areas (e.g., agricultural science), the vast majority are liberal arts degrees. Conversely, associate of applied science (AAS) degrees are terminal and apply to specific, commonly technical fields. Women who enroll in college following entry but do not earn a degree spend less time in school and only earn 19 credits, on average.

Not surprisingly, in the initial quarters after entry, college-going women earn less, on average, than women who do not attend college. But four years following entry, college

[7] Common examples of certificates awarded to welfare recipients include real estate, computer information systems, emergency medical services, and nurse assistant/home health aid.

attendees earn approximately $400 more per quarter than their counterparts who do not attend college, while degree recipients earn close to $1000 more.

2.4.2 Graphical Analysis

Figures 2.2 and 2.3 preview my main approach and results. Figure 2.2 graphs unadjusted average earnings by quarter, both before and following welfare entry. The black solid line indicates individuals who attend college following entry into Colorado Works but do not earn a credential following entry (drop-outs), the thick dashed line represents the earnings trajectories of degree recipients, and the gray line represents the earnings of women who do not attend college prior to or following welfare entry. I exclude women who were already enrolled in college at entry (described in the fourth column of Table 2.1). Thin dotted lines represent 95 percent confidence intervals.

Figure 2.2 shows an "Ashenfelter dip" in the quarters immediately surrounding entry. All sample members experience a decline in earnings in the two to three quarters before and following welfare entry. Up to three years after welfare entry, degree recipients have lower earnings than other women, likely due to a greater number of quarters spent in college. Conversely, drop-outs appear to have much lower foregone earnings, but their earnings gains following welfare entry are small. Beginning in the 12th quarter after entry, the earnings of degree recipients significantly exceed both the earnings of other groups and their own pre-welfare earnings, although these gains are imprecisely estimated due to the small group size.

Figure 2.3 replicates this exercise, graphing residual earnings from a regression on individual fixed effects. The dashed line without a confidence interval represents the average quarterly county unemployment rate. The largest earnings gains to degree recipients accrue at the end of the sample period, when unemployment is increasing. During this same period, the residual earnings of college drop-outs and women who do not attend college are indistinguishable and decline to below pre-entry levels at approximately 18 quarters after entry.

2.5 Event Study Framework

Simple regressions of wages on college enrollment will be biased if the decision to attend college is correlated with unobservable individual characteristics that affect employment and earnings. Selection bias is likely a concern even among welfare recipients; London (2006) shows that pre-PRWORA women attending college while on welfare had higher test scores and were more likely to have a two-year institution in their county of residence.

To address this concern, I follow an approach similar to Jacobson et al. (2005), taking advantage of the fact that I observe earnings for women both before and after welfare receipt and college attendance. If selection bias is only driven by time-invariant unobservable characteristics, then including an individual fixed effect will deal with the endogeneity of college attendance and degree receipt. Essentially, this approach differences out a person-specific mean level of wages from observed quarterly wages, using observations of individuals' wages prior to college attendance as a counterfactual for current wages in the absence of college-going. Additionally, I include observations of individuals who do not attend college to identify year and quarter fixed effects and the impacts of other observable characteristics. The key identifying assumption is that there are no time-varying unobservable shocks correlated with education and labor market outcomes affect college-going women differentially than other recipients.

I estimate the following linear regression model:

$$(1) \qquad y_{it} = \alpha_i + \lambda f(c_{it}, d_{it}) + \delta e_{it} + \beta X_{it} + \eta_{year} + \eta_{quarter} + \eta + t_{pre}\eta_{cohort} + t_{post}\eta_{cohort} + \varepsilon_{it}$$

where y_{it} is the outcome of interest (i.e., quarterly earnings, probability of employment) for individual i in quarter t, α_i is an individual-specific fixed effect that controls for time-invariant unobservable qualities, such as ability or motivation, e_{it} is an indicator for whether an individual is enrolled in college in quarter t, since individuals may substitute hours of employment for hours of school while enrolled, and X_{it} is a vector of time-varying observable individual characteristics.[8] I include year, quarter and county fixed effects and a entry cohort specific time trend, allowing for a trend break at welfare entry (e.g., $t_{pre}\eta_{cohort}$ is the pre-entry trend); ε_{it} is an individual-specific error term that is uncorrelated with educational attainment.

Finally, $f(c_{it}, d_{it})$ is a function of credits and credentials received at the end of college enrollment; λ identifies the causal effect of education if there are no time varying unobservable characteristics correlated both with college attendance and employment/welfare outcomes. I examine two specifications of this function. My first specification includes a linear term in credits, ignoring any credentials received. In my second specification, I add indicators for degree receipt and again include a linear term in credits received interacted with an indicator for not having earned a credential: $\lambda f(c_{it}, d_{it}) = \lambda_1 c_{it} 1[d \in \{D^d\} = 0] + \sum_d \lambda^d D^d$, where D^d is a vector of indicator variables for receipt of specific degrees (i.e., certificates, AA/AGS, AAS). This specification allows me to test the impacts of college-going on labor-market outcomes for women who ultimately drop-out.

2.5.1 Estimates of the Impacts of College Credits and Credentials on Labor Market Outcomes

Table 2.2 examines the effect of college attendance and number of credits received on wages and employment. Columns 1, 3, 5, and 7 are cross-sectional estimates, while the

---

[8] These characteristics include a quadratic term in age, number of children, age of youngest child, vehicle ownership, and indicators for months of lifetime welfare receipt (0 months, 1 to 12, 13 to 24, 25 to 59, and 60 or more months). These intervals correspond to cut-offs that trigger changes in Colorado welfare rules. For instance, two thirds of a participant's income is disregarded from benefit calculation for the first twelve months of participation. After 24 months of assistance, recipients are required to participate in at least 20 hours of work related activities per week (e.g., employment, job search, on the job training). Finally, Colorado follows the federal 60 month lifetime limit on benefit receipt (although some participants are granted extensions for extenuating circumstances).

remaining columns include individual fixed effects. Quarterly earnings and the probability of employment are both increasing in the number of credits earned when I do not account for credentials. In the cross-section, the impact of credits on earnings and employment remains positive and significant, suggesting that one year of full-time college attendance leads to a $240 increase in quarterly earnings (an approximately 14 percent increase from pre-welfare earnings) and a 3 percentage point increase in the probability of being employed in a given quarter. However, when I include individual fixed effects, credits are no longer significantly related to earnings or employment. Surprisingly, the coefficients representing the impacts of degree receipt are quite similar in the cross sectional and fixed effects estimates.

All vocational degrees and credentials offered by community colleges in Colorado significantly raise quarterly earnings and employment. Even a short-term certificate, requiring less than one year of full-time attendance, increases earnings by $480 per quarter (a 28 percent increase from pre-welfare earnings). Longer term certificates result in slightly higher (but not statistically distinguishable) gains. Women who earn an AAS degree experience the largest increases in earnings and employment among community college attendees, with their earnings increasing by close to $2,500 per quarter (over a 100 percent increase). BA degree recipients see similar earnings gains, but the number of such individuals in my sample is quite small, so estimates relating to the outcomes of four-year college graduates should be interpreted with caution. Academic and general studies associate's degree recipients do not experience any earnings or employment gains once time-invariant characteristics are accounted for.

Table 2.3 subjects my main specification to a variety of robustness tests. In the first and fifth columns, I only examine women with a strong pre-welfare labor market attachment, to eliminate concerns that my results are driven by a negative correlation between employment and post-entry college attendance. In columns 2 and 6, I eliminate women who were already attending college before entering welfare. In columns 3 and 7, I eliminate individuals who are still attending college at the end of my sample period, and in the last specification, I drop the quarters surrounding welfare entry where earnings display a dip. My results are quite consistent across specifications, with credentials and vocational associate's degrees strongly correlated with earnings and employment gains, while AA/AGS degrees and credits alone do not result in improved labor market outcomes.

In Table 2.4, I compare the impacts of credits and credentials, distinguishing math, science, and health-focused programs from other programs offered by community colleges. Jacobson et al. (2005) find that displaced workers who earn credits in these fields benefit the most on the labor market. Results including all credentials and credits by program are displayed in columns 1 and 3. I do not find evidence that credits from health, math, or science programs increase earnings or employment. Longer-term health, math, and science credentials lead to larger earnings gains than credentials from other programs, although the impacts of short term certificates on earnings and employment do not vary significantly by program. However, Jacobson et al. (2005) only observe associate degree receipt and do not observe whether individuals in their sample earn certificates. Thus, in columns 2 and 4, I treat certificates as unobservable and do find evidence of a positive impact of credits from math, health, and science programs on earnings (although not employment). When certificates are unobservable, estimates would erroneously suggest that credits from one year of full-time attendance in these programs leads to a $135 increase in quarterly earnings (approximately an 8 percent gain over pre-welfare earnings).

2.5.2 Short-run Private and Social Returns

Despite the large estimated impacts of credential receipt on employment and earnings, only 10 percent of women who attend college after entering Colorado's welfare program obtain a certificate or degree. It is difficult to determine whether this behavior is optimal (e.g., women learn about their "type" after enrolling in college and realize their own personal returns are too small), due to a market failure (e.g., credit constraints), or a result of self-control issues (e.g., hyperbolic discounting). Without a source of exogenous variation in these parameters, it is difficult to determine which, if any, affect completion rates.

Nonetheless, in Table 2.5, I document the extent to which credit constraints may be an issue for women in this population. My estimates thus far have looked at the post-enrollment impacts of college-going on labor market outcomes, not accounting for foregone earnings while attending college. Additionally, from the Colorado Department of Higher Education, I observe financial aid women received while in college and the total cost of attendance in a given semester, giving me an accurate measure of the direct costs of college-going. The first two columns of Table 2.5 take these indirect costs into consideration. The dependent variable in column 1 is quarterly earnings, but I account for the indirect cost of schooling due to foregone earnings by including the impacts of degree and credit receipt both before and after college exit. Since I only observe women who earn a credential for one or two years after exit, while, as shown in Figure 2.2, women spend up to three years in college, these estimates will represent short-run returns. In comparison, the estimates presented in the beginning of this section represent medium- to long-run returns (depending on assumptions about age-earnings profiles).

After accounting for foregone earnings, the only community college credential recipients who experience a positive return on their investment in the short-run are women who earn AAS degrees, and the estimated impact on quarterly earnings falls to $500 per quarter. After accounting for direct and indirect costs of college attendance, even these women do not experience a significant return on their investment. Estimated returns to all other community college credentials and credits are negative and in some cases quite large. For instance, women who earn a longer-term certificate experience a $1,200 reduction in quarterly earnings. Although the majority of women who attend college receive both grants and loans from Colorado and the federal government, tuition and fees are substantially higher than the value of this aid.[9]

In the third and fourth columns of Table 2.5, I investigate the extent to which Colorado's safety net can mitigate these costs by taking into account the value of cash assistance (column 3) and subsidized child care (column 4) while on welfare.[10] Cash assistance provide through Colorado's welfare program does little to cushion women from the costs of college attendance. Even after accounting for the potential value of child care subsidies, most credential recipients do not experience a positive rate of return to their investment in the short run. If many of my sample members are credit constrained, in the short-run these negative returns could reduce educational attainment.

In the final two columns of Table 2.5, I attempt to investigate the potential social returns to college attendance by estimating impacts of credit and degree receipt on the probability of reentering welfare and the value of cash assistance received after a sample member's initial spell. I find little evidence in the way of positive social returns to higher education. However, the latter part of my sample period also coincides with changes in federal policy that increased the

[9] I treat loans as grant aid while in school and assume 30 years of repayment following college exit.

[10] Women on welfare with children under 13 are eligible for child care subsidies. I do not directly observe child care subsidies, which vary by county and child age. However, in each county, the reimbursement rate is set to 75 percent of the market rate, which I do observe and use to impute the value of a mother's potential subsidy based on the number and age of her children.

pressure on states to reduce caseloads and enforce participation requirements for current recipients, making these results suggestive at best.

2.6 Conclusions

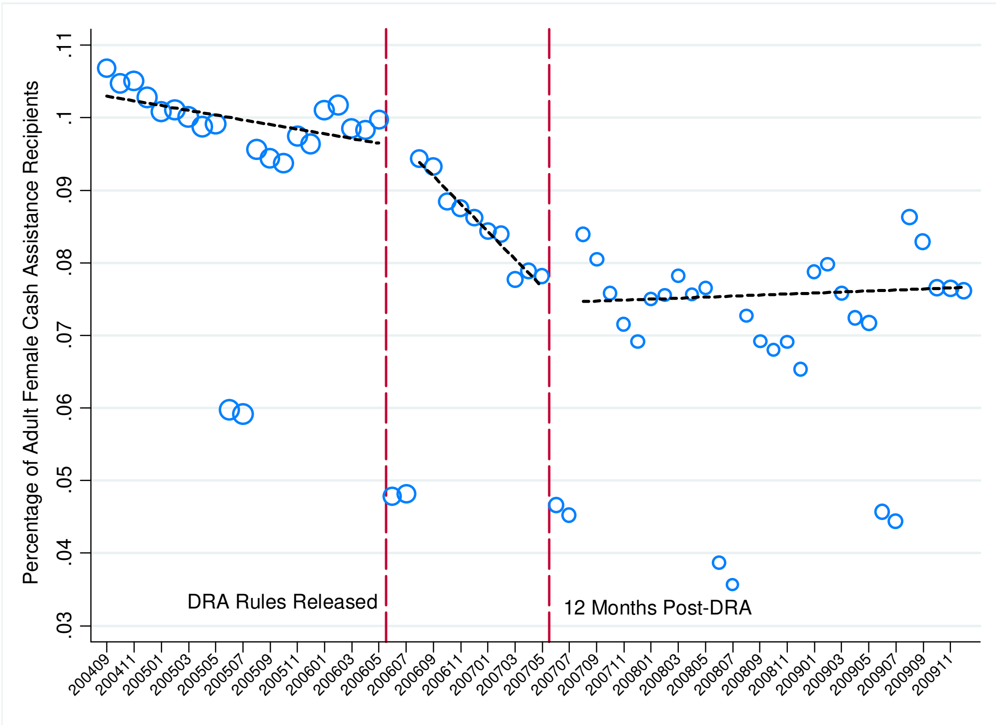
Using an individual fixed effects approach, I find that women at risk for welfare receipt benefit from college attendance, although the impacts driven credential receipt. Even women who receive a certificate for less than one year of full-time study see their earnings increase by close to $500 per quarter. These are significant gains, given that on average, these women only earned $1,700 per quarter prior to welfare entry. Women who earn a terminal associate's degree see their earnings increase by close to $2,500 per quarter. The sole exception is for women receiving associate of arts and general studies degrees. These women experienced only small increases in earnings. Credentials earned in health, math, and science programs lead to larger earnings gains. However, converse to prior findings, credits earned in these programs do not appear to affect earnings or employment once certificate receipt is accounted for.

My results and approach have important limitations. Since I only observe individuals for up to 14 quarters after welfare entry, I can only estimate the medium-term impacts of credentials on labor market outcomes. If some degrees, such as AA/AGS, take longer to produce earnings gains, results will not represent the long-run impact of higher education. Second, including individual-specific fixed effects essentially estimates individual-deviations from mean wages after college attendance and degree receipt. However, if time-varying unobservable characteristics are correlated with both the decision to attend college and outcomes, my estimates will still include this selection bias. Finally, I only observe college attendance and degree receipt at public institutions in Colorado. Cellini and Chaudhary (2011) find that the returns to private, two-year degrees are similar to the returns to degrees earned at community colleges. It is likely that some women in my sample attended one of the 312 private technical schools licensed to operate in Colorado. Thus, my estimates are likely a lower bound of the returns to higher education.

These results have implications for how welfare policy and policy aimed at two-year schools can positively impact the labor market outcomes for low-skilled individuals. In particular, these results suggest that providing supports to ensure community college students receive a degree, even certificates for short-term periods of attendance, can substantially increase earnings. Many recent policies aimed at potential community college students primarily focus on college entry. My results suggest that the marginal student affected by these policies may not benefit substantially unless additional supports for degree completion are provided.

2.7 Figures and Tables

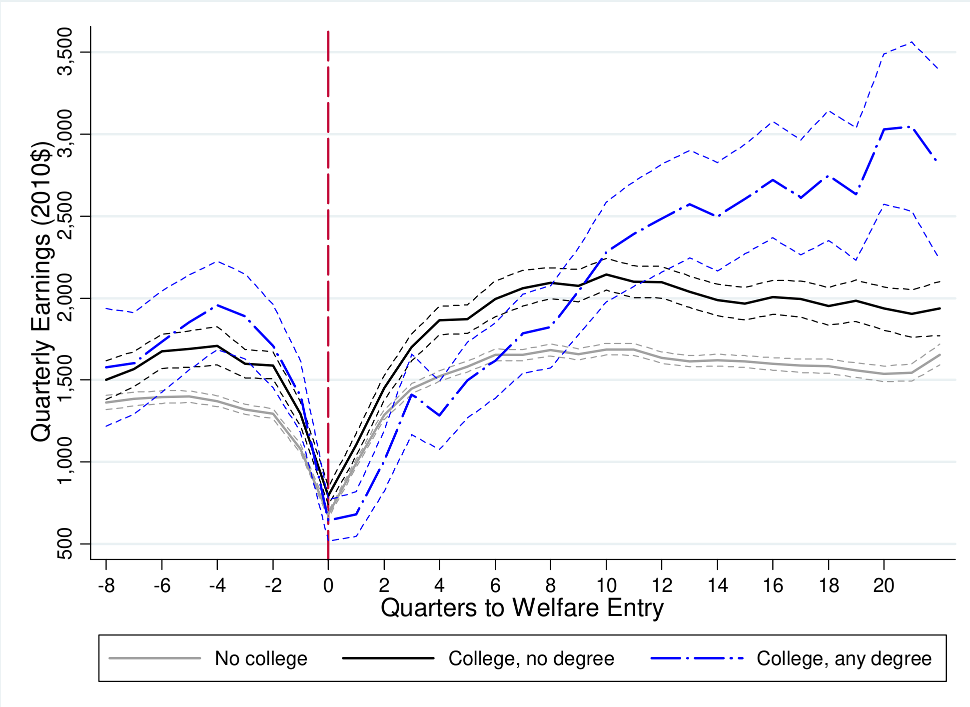
Figure 2.1: Concurrent College Attendance and Cash Assistance Receipt, September 2004 – December 2012



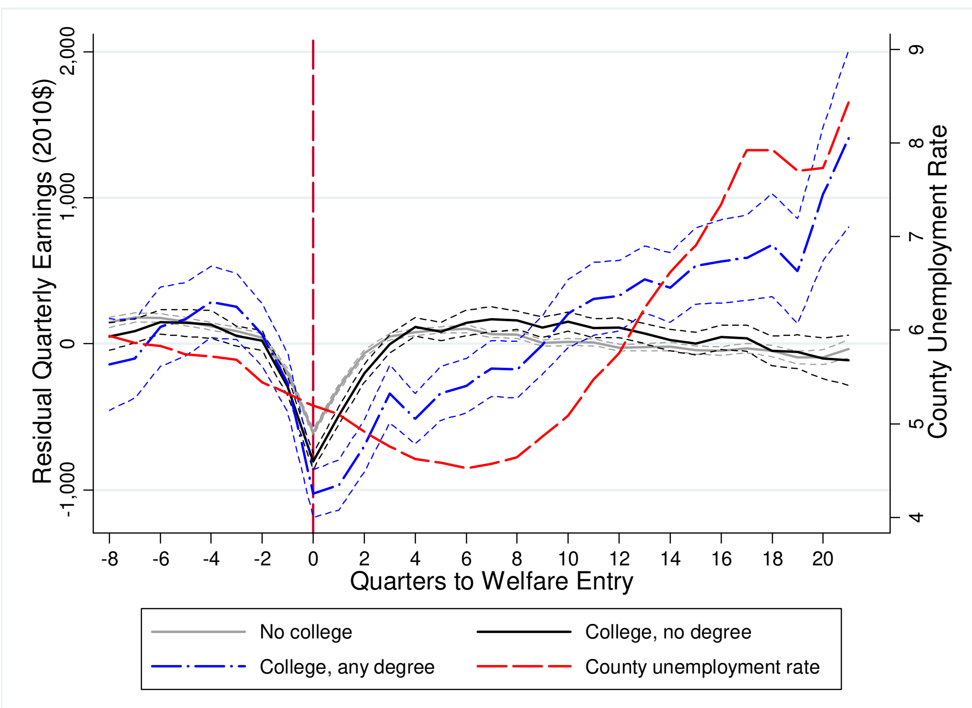
Notes: Each circle represents the percentage of adult female TANF recipients attending a public college, with larger circles indicating larger total caseloads. The black dashed lines represent a linear fit of college going for three periods: before the Deficit Reduction Act (DRA) final rules were released, between 0 and 12 months after the rules were released, and 12 or more months after the rules were released.

Figure 2.2: Average Quarterly Earnings by Educational Attainment following Welfare Entry

A. Raw Means



B. Residual



Notes: Thick lines represent average quarterly earnings before and after welfare entry by educational attainment and credential receipt. Thin dashed lines represent 95 percent confidence intervals.

Table 2.1: Characteristics of Colorado Works Recipients by College Attendance and Degree Receipt

	No College	College Attendance		In College Pre-Entry
		No Degree	Any Degee	
Number of individuals	23,115	3,329	380	2,732
Race				
Hispanic	0.31	0.28	0.32	0.28
Native American or Asian/Pacific Islander	0.04	0.04	0.03	0.04
Black	0.17	0.21	0.12	0.2
White	0.49	0.47	0.53	0.48
Characteristics at welfare entry				
Age	30	28	28	28
Never married	0.53	0.59	0.58	0.60
Number of children	2.0	1.9	1.8	1.9
Age of youngest child	4	4	3	4
Lifetime months on welfare	9.2	8.9	6.0	11.0
First spell on welfare	0.20	0.22	0.23	0.16
Own vehicle	0.29	0.34	0.39	0.40
Disabled	0.13	0.10	0.08	0.12
Average quarterly earnings before entry (2010$)	$1,357	$1,642	$1,768	$1,544
Percentage of quarters worked before entry	0.41	0.49	0.52	0.51
Prior college attendance (1990 or later)	0.15	0.27	0.34	1.00
Prior credits earned (1990 or later)	4	10	19	50
Outcomes				
Percentage of quarters employed after entry	0.40	0.51	0.53	0.51
Average quarterly earnings (2010$)				
All quarters after entry	$1,589	$1,870	$2,092	$2,119
1 year after entry	$1,549	$1,890	$1,293	$1,839
2 years after entry	$1,707	$2,139	$1,854	$2,298
3 years after entry	$1,654	$2,133	$2,541	$2,510
4 years after entry	$1,614	$2,017	$2,765	$2,456
Percentage of quarters on welfare	0.25	0.29	0.32	0.28
Percentage of quarters in school	0	0.24	0.42	0.23
Credits earned	0	19	67	25
Degree received				
Certificate (less than 1 year)	0	0	0.53	0.03
Certificate (1 to less than 2 years)	0	0	0.14	0.02
Associate of Arts/General Studies	0	0	0.17	0.05
Associate of Applied Science	0	0	0.11	0.02
Bachelor's Degree	0	0	0.05	0.04

Data: CDHS program data, CDLE quarterly earnings records, and CDHE enrollment and degree files. **Notes:** Sample includes female adult Colorado Works recipients who began a spell of welfare receipt between 10/2004 and 6/2007. Categories based on college-going after entering welfare. Race and marital status for women with non-missing values.

Table 2.2: Cross-Sectional & Event Study Estimates of the Impact of College Credits & Degree Receipt on Employment & Earnings

	Dependent Variable = Quarterly Earnings				Dependent Variable = Pr(Employed)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Credits	7.95	5.91	7.84	1.02	0.001	0.0003	0.001	0.0001
	(0.54)**	(0.69)**	(0.79)**	(0.65)	(0.000)**	(0.0001)**	(0.000)**	(0.0001)
*Point estimate: credits*30*	*238.6*	*177.4*	*235.3*	*30.71*	*0.024*	*0.010*	*0.034*	*0.003*
	*(16.18)***	*(20.67)***	*(23.66)***	*(19.41)*	*(0.002)***	*(0.002)***	*(0.003)***	*(0.003)*
Degree:								
Short-term Certificate			417.80	476.91			0.089	0.098
			(87.80)**	(120.54)**			(0.013)**	(0.021)**
Certificate			732.76	682.32			0.089	0.123
			(150.63)**	(223.84)**			(0.020)**	(0.028)**
Associate of Applied Science			1,450.50	2,415.04			0.140	0.180
			(194.90)**	(368.80)**			(0.019)**	(0.031)**
Associate of Arts/GS			267.16	-152.19			0.048	0.020
			(180.88)	(181.28)			(0.022)*	(0.034)
Bachelor's Degree			2,125.76	2,613.22			0.126	0.199
			(258.02)**	(292.69)**			(0.019)**	(0.033)**
Observations	857,124	857,124	857,124	857,124	857,124	857,124	857,124	857,124
Individual fixed effects		X		X		X		X

Data: CDHS program data, CDLE quarterly earnings records, CDHE enrollment and degree files. **Notes:** + significant at 10%; * significant at 5%; ** significant at 1%; robust standard errors clustered by individual in parentheses; each column denotes a separate regression. Regressions include year and quarter fixed effects, an indicator for college enrollment, and a linear cohort trend allowing for a trend break at entry. Regressions also control for vehicle ownership, number of children, age of youngest child, presence of a disability, months on cash assistance (0, 1 to 12, 13 to 24, 25 or more), quarterly county unemployment rates, and a quadratic in age. Column 1, 3, 5, and 7 also control for race. All dollar amounts adjusted for inflation (2010$).

Table 2.3: Robustness of the Impact of Credits and Degrees on Labor Market Outcomes to Alternative Samples

	Dependent Variable = Quarterly Earnings				Dependent Variable = Pr(Employed)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Credits	0.98	0.90	0.97	0.43	0.0001	0.0003	0.0001	0.000004
	(0.85)	(0.93)	(0.67)	(0.67)	(0.0001)	(0.0002)+	(0.0001)	(0.0001)
*Point estimate: credits*30*	*29.47*	*26.95*	*29.15*	*12.82*	*0.002*	*0.009*	*0.003*	*0.0001*
	(25.42)	*(27.99)*	*(20.17)*	*(20.01)*	*(0.004)*	*(0.005)+*	*(0.003)*	*(0.003)*
Degree:								
Short-term Certificate	475.09	710.74	473.58	501.11	0.107	0.131	0.098	0.089
	(136.43)**	(153.94)**	(122.71)**	(139.64)**	(0.022)**	(0.026)**	(0.021)**	(0.024)**
Certificate	762.21	587.79	647.03	613.15	0.147	0.111	0.115	0.116
	(298.66)*	(251.81)*	(230.52)**	(265.49)*	(0.033)**	(0.045)*	(0.028)**	(0.030)**
Associate of Applied Science	2,594.59	2,099.75	2,452.28	2277.59	0.209	0.230	0.180	0.179
	(444.82)**	(491.40)**	(378.24)**	(352.41)**	(0.033)**	(0.048)**	(0.031)**	(0.030)**
Associate of Arts/GS	-346.46	-162.60	-188.40	-99.35	0.020	0.020	0.023	0.009
	(200.54)+	(318.24)	(185.49)	(198.16)	(0.041)	(0.061)	(0.035)	(0.038)
Bachelor's Degree	2,511.35	3,035.85	2,642.84	2,738.86	0.132	0.276	0.195	0.196
	(402.70)**	(751.00)**	(302.48)**	(317.22)**	(0.037)**	(0.091)**	(0.033)**	(0.036)**
Observations	525,886	777,896	813,653	620,676	525,886	777,896	813,653	620,676
Strong LF attachment	X				X			
Not in school pre-entry		X				X		
Not in school at end of sample			X				X	
4+ quarters from entry				X				X

Data: CDHS program data, CDLE quarterly earnings records, and CDHE enrollment files. **Notes:** + significant at 10%; * significant at 5%; ** significant at 1%; robust standard errors clustered by individual in parentheses; each column denotes a separate regression. Regressions include year, quarter, and individual fixed effects, an indicator for college attendance, and a linear cohort trend allowing for a trend break at entry. Regressions also control for vehicle ownership, number of children, age of youngest child, presence of a disability, months on cash assistance (0, 1 to 12, 13 to 24, 25 or more), quarterly county unemployment rates, and a quadratic in age. Regressions displayed in columns 1 and 5 exclude women who were employed for less than 50 percent of the quarters before welfare entry, columns 2 and 6 exclude women who were enrolled in one of the three quarters prior to entry, columns 3 and 7 exclude women still enrolled in school during the last quarter observed, and columns 4 and 8 observations between four quarters before entry and four quarters after entry (eliminating quarters in the "Ashenfelter dip"). All dollar amounts adjusted for inflation (2010$).

Table 2.4: The Returns to College Credits and Degrees by Program

	1. Quarterly Earnings		2. Pr(Employed)	
	(1)	(2)	(3)	(4)
Credits | No degree				
Health/Math/Science	1.84	4.53	0.0002	0.0002
	(1.83)	(1.91)*	(0.0004)	(0.0003)
Other program	0.29	0.03	0.0001	0.0001
	(0.96)	(0.81)	(0.0002)	(0.0001)
Test of equality (p-value)	*0.484*	*0.043*	*0.996*	*0.866*
Short-term Certificate				
Health/Math/Science	619.14	--	0.113	--
	(153.60)**		(0.028)**	
Other	374.09	--	0.060	--
	(179.27)*		(0.032)+	
Test of equality (p-value)	*0.313*		*0.210*	
Certificate				
Health/Math/Science	1,538.21	--	0.159	--
	(406.24)**		(0.042)**	
Other two-year program	415.82	--	0.128	--
	(183.87)*		(0.036)**	
Test of equality (p-value)	*0.013*		*0.586*	
Associate of Applied Science				
Health/Math/Science	3,904.54	3,674.91	0.241	0.213
	(569.33)**	(563.18)**	(0.042)**	(0.041)**
Other two-year program	439.45	416.94	0.084	0.077
	(335.82)	(333.20)	(0.044)+	(0.044)+
Test of equality (p-value)	*0.000*	*0.000*	*0.011*	*0.025*
Associate of Arts/GS	-156.31	-141.91	0.026	0.025
	(186.29)	(185.66)	(0.034)	(0.034)
Bachelor's Degree	2,646.35	2,646.82	0.201	0.200
	(294.02)**	(293.61)**	(0.033)**	(0.033)**
Observations	857,124	857,124	857,124	857,124

Data: CDHS program data, CDLE quarterly earnings records, and CDHE enrollment and degree files. **Notes:** + significant at 10%; * significant at 5%; ** significant at 1%; robust standard errors clustered by individual in parentheses; each column denotes a separate regression. Regressions include year, quarter, and individual fixed effects, an indicator for college attendance, and a linear cohort trend allowing for a trend break at entry. Regressions also control for vehicle ownership, number of children, age of youngest child, presence of a disability, months on cash assistance (0, 1 to 12, 13 to 24, 25 or more), quarterly county unemployment rates, and a quadratic in age. Columns 2 and 4 assume certificates are unobservable. All dollar amounts adjusted for inflation (2010$).

Table 2.5: Short-Run Private and Social Returns to College Credits and Degrees

	Private Returns, Accounting for:				Social Returns	
	1.Foregone earnings	2. Direct costs	3. Cash assistance	4. Child care subsidy	5. Pr(Welfare)	6. Total Cash Assistance
Credits	-5.18	-9.47	-8.76	-7.54	-0.0004	-0.49
	(0.74)**	(0.79)**	(0.78)**	(0.85)**	(0.0002)*	(0.25)+
*Point estimate: credits*30*	*-155.3*	*-284.2*	*-262.7*	*-226.3*	*-0.013*	*-14.82*
	*(22.13)***	*(23.7)***	*(23.31)***	*(25.5)***	*(0.005)**	*(7.63)+*
Degree:						
Short-term Certificate	21.95	-363.18	-273.85	-47.11	-0.027	-9.97
	(147.47)	(151.42)*	(148.32)+	(160.62)	(0.024)	(42.68)
Certificate	-1,008.22	-1,259.01	-1,163.56	-1,043.60	-0.006	49.62
	(413.49)*	(439.29)**	(424.83)**	(415.93)*	(0.032)	(60.01)
Associate of Applied Science	514.33	355.92	467.59	597.80	0.016	42.27
	(305.88)+	(339.41)	(339.06)	(378.21)	(0.045)	(87.46)
Associate of Arts/GS	-1,010.13	-1,546.15	-1,461.16	-1,161.53	-0.070	-79.81
	(233.38)**	(233.32)**	(231.74)**	(244.66)**	(0.042)+	(53.04)
Bachelor's Degree	793.20	805.54	891.98	1,126.37	-0.200	-161.19
	(328.77)*	(351.51)*	(351.78)*	(356.37)**	(0.055)**	(69.20)*
Observations	827,568	827,568	827,568	827,568	524,231	524,231

Data: CDHS program data, CDLE quarterly earnings records, and CDHE enrollment, degree, and financial aid files. **Notes:** + significant at 10%; * significant at 5%; ** significant at 1%; robust standard errors clustered by individual in parentheses; each column denotes a separate regression. Regressions include year, quarter, and individual fixed effects, an indicator for college enrollment, and a linear cohort trend allowing for a trend break at entry. Regressions also control for vehicle ownership, number of children, age of youngest child, presence of a disability, months on cash assistance (0, 1 to 12, 13 to 24, 25 or more), and a quadratic in age. All dollar amounts adjusted for inflation (2010$). Columns 1 through 4 account for foregone earnings while attending school (see text). Dependent variables are as follows: quarterly earnings (column 1), quarterly earnings less direct costs of college attendance (column 2), quarterly earnings and cash assistance less direct costs (column 3), quarterly earnings, cash assistance, and the value of a mother's potential child care subsidy less direct costs (column 5), welfare receipt after initial entry (column 5), and cash assistance after entry (column 6).

CHAPTER 3

The Design of Teacher Incentive Pay and Educational Outcomes: Evidence from the New York City Bonus Program

3.1 Introduction

Teacher compensation schemes are often criticized for their lack of performance pay. In other sectors, incentive pay increases worker effort and output by aligning the interests of workers and employers, providing information about the most valued aspects of an employee's job, and motivating workers to provide costly effort (Gibbons 1998; Lazear and Oyer 2010). In this paper, we examine a group-based teacher incentive scheme implemented by the New York City Department of Education (DOE) and investigate whether specific features of the program contributed to its ineffectiveness.

In 2007, close to two hundred schools were randomly selected from a group of high-poverty schools.[1] These schools could earn school-wide bonuses by surpassing goals primarily based on student achievement. Successful schools would earn lump sum payments equal to \$3000 per union teacher (three to seven percent of annual teacher pay). Several independent studies show that the bonus program had little overall effect on either math or reading achievement (Springer and Winters 2009; Goodman and Turner 2010; Fryer 2011). We show that in schools where smaller groups of teachers were responsible for instructing tested students, the program led to small but significant increases in student achievement. Our finding is consistent with predictions that group-based incentives are diluted by the potential for free-riding when payments depend on actions of a large number of workers (Holmstrom 1982).

Several features of the educational sector complicate the design of teacher performance pay. First, performance pay is most effective when employers can measure worker output or when observable effort and productivity are closely aligned. Monitoring teachers is costly and

[1] This experiment was designed and implemented by the New York City Department of Education and teachers' union, random assignment was conducted by Roland Fryer, and RAND performed the official evaluation.

measuring individual teachers' contributions to student achievement is difficult. Second, although education is a complex good and teachers must allocate their effort across several activities, teacher incentive pay is often linked to a single performance measure (e.g., student test scores), which may lead teachers to direct effort away from other beneficial classroom activities (Holmstrom and Milgrom 1991).[2] Despite these issues, studies from outside the United States demonstrate that teacher incentive pay can increase student achievement (e.g., Lavy 2002; Lavy 2009; Muralidharan and Sundararaman 2011).

Specific features of the NYC bonus program may have limited its effectiveness. First, the program linked incentive pay to school-wide performance goals**.** In theory, group incentive pay is most effective in the context of a joint production technology (Itoh, 1991). For instance, if an individual teacher's effort has positive impacts on the effort exerted by her peers (e.g., Jackson and Bruegmann 2009), group incentives may outperform individual incentives. Otherwise, relative to individual incentives, group incentives decrease individual returns to effort and will lead to free-riding unless workers monitor each other's effort.

We test for free-riding by allowing the bonus program's impacts to vary by the number of teachers with students who are tested (and therefore contribute to the probability that a school qualifies for the bonus award). To test for the importance of joint production and monitoring, we examine whether program impacts vary by the degree to which teachers report collaborating in lesson planning and instruction using a survey administered prior to program implementation. We show that the bonus program raised math achievement in schools with a small number of teachers with tested students, although these impacts are small (0.08 student-level standard

[2] Teachers may also be induced to focus on narrow, exam-related basic skills, manipulate test scores, or focus on students whose performance contributes more towards goals (e.g., Jacob and Levitt 2003; Jacob 2005; Cullen and Reback 2006; Neal and Schanzenbach 2010).

deviations) and only marginally significant in the program's second year. We present suggestive evidence of positive program impacts in schools with a high degree of collaboration.

Second, teachers already faced negative incentives when the bonus program was implemented. In fall 2007, the DOE instituted a district-wide accountability system that imposed sanctions on schools that did not meet the same goals used in determining bonus receipt. Thus, estimated impacts of the bonus program represent the effect of teacher performance pay in schools already under accountability pressure. However, this may be the most appropriate context to examine, since many states have implemented accountability systems and all public school districts face pressure from No Child Left Behind provisions. Finally, we find no differences in the impacts of the bonus program when we compare schools under different degrees of accountability pressure, suggesting that our results are not solely driven by the dilution of incentives due to the accountability system.

Third, teachers' lack of understanding of the bonus program's complex goals may have limited its efficacy. Alternatively, since bonus awards were provided if a school's performance reached a set threshold, if thresholds were set too high or too low, a large number of teachers may have optimally responded by not changing their behavior (Neal 2011). However, the metrics used to determine bonus payments were the same goals used by the district-wide accountability system and Rockoff and Turner (2010) show that negative incentives provided through this system increased student achievement.[3]

The next section of this chapter describes New York City's bonus program. Section 3.3 provides an overview of the data and estimation framework while Section 3.4 presents results and the fifth section concludes.

---

[3] On a related note, a committee within each school had some discretion over how bonuses would be distributed. However, the distribution scheme was set *ex ante* and most schools chose equal or close to equal distributions.

3.2 The New York City Bonus Program

We use a policy experiment implemented by the New York City Department of Education (DOE) in the fall of 2007, the "School-Wide Performance Bonus Program" (hereafter, the bonus program). Both the DOE and the United Federation of Teachers (UFT) endorsed the program as an innovative model for teacher performance pay. In November 2007, 181 schools serving kindergarten through eighth grade were randomly selected from a group of 309 schools designated as "high need"; 128 schools were assigned to the treatment group. Two of the 181 schools originally assigned to the treatment group were moved to the control group prior to notification of their assignment; for the purposes of our analyses, we consider these schools as part of the original treatment group. Treatment schools were eligible to participate in the program, contingent on 55 percent of a school's full-time United Federal of Teachers (UFT) staff voting in favor of participation. Twenty-five schools voted not to participate or withdrew from the program following a vote of approval. Finally, four schools originally assigned to the control group were allowed to vote and ultimately chose to participate in the bonus program; for the purposes of our analyses, we consider these schools as part of the original control group. However, the group of schools that ultimately could earn bonus payments totaled 158.

The schools that voted in favor of the program could earn a lump-sum bonus by meeting a school-wide goal. These goals were tied to the NYC accountability system which awarded letter grades to schools (explained below) and were primarily based on student achievement on state math and reading exams. Schools that achieved their goals received lump sum bonuses equal to $3,000 per union teacher, while schools that fell short but managed to meet 75 percent of their goal received $1,500 per union teacher. Thus, although total bonus awards varied across schools with different numbers of union teachers, the expected bonus payment was equal across these schools. Schools that did not reach their target suffered no consequences beyond the absence of bonus pay. The full $3,000 award represents a seven percent increase in the salary of teachers at the bottom of the pay scale and a three percent increase for the most experienced teachers.[4]

Each participating school selected a four-member compensation committee, consisting of the principal, a second administrator, and two union representatives elected by the school's UFT members. In the program's first year, this committee was required to submit a bonus distribution scheme after students took the state math and reading exams but before exam results were released. Thus, at least in the first year of the program, teachers' effort decisions should not be affected by the distribution that was ultimately chosen. Program guidelines stipulated that within schools reaching their goal, all union teachers must receive a bonus payment and individual bonuses could not be explicitly based on seniority. Beyond these requirements, committees had complete freedom in determining individual teachers' bonus payments and could also provide bonus payments to other school employees. Around half of treatment schools chose an approximately equal distribution (i.e., the difference between the highest and lowest bonus payment was less than $100). In the remainder of schools, the difference between the highest and lowest bonus ranged from $200 to $5000.

The 2007-2008 school year also marked the implementation of the DOE's new accountability system. Under this system, schools received accountability grades designed to summarize a school's overall performance on a multidimensional metric of student learning.[5]

---

[4] Similar to the majority of public school districts in the U.S., teacher salaries in New York City are determined through a schedule based on years of experience and graduate coursework (Podgursky and Springer 2007).

[5] The metric includes a measure of school environment (student attendance and results from survey of parents, teachers, and students), student performance (average student achievement on reading and math exams, median

Each school's performance was scored relative to the entire district and to a group of peer schools. This group included the 40 schools that were most similar according to a "peer index" that was based on student demographic characteristics and prior achievement.[6] Each school received a progress report documenting its overall performance, the corresponding accountability grade, and a target score for the following year. Schools with lower accountability grades needed to make larger improvements to reach their targets. Importantly, these target scores determined which schools participating in the bonus program received awards.

Moreover, the accountability system provided additional incentives to improve student achievement, regardless of bonus program participation. Schools that earned an A or B accountability grade received rewards (e.g., principal bonuses, additional funds when students transferred from schools receiving a poor grade), while schools that received D and F grades faced consequences (e.g., school closure and principal removal). Although this accountability system was more complex than systems based on a single metric (e.g., the percentage of students achieving proficiency), teachers and administrators received training on how to interpret the complicated set of measures determining a school's grade, and it was clear that grades were largely determined by student performance on math and reading exams. Rockoff and Turner (2010) find that receiving an F or D led to a significant improvement in student test scores, a result consistent with school employees understanding that performance under the accountability system was dependent on student achievement. Bonus program impacts do not vary across schools with different accountability grades (see Section 4). However, it is still important to note

---

proficiency, and percentage students achieving proficiency), and student progress (average change and percent making progress on math and reading exams). Schools received extra credit for progress among high-need students.

[6] For elementary schools and schools serving kindergarten through eighth grade (K-8), the peer index was based on a function of the percentage of students that were English language learner (ELL), special education, Title I free lunch, and minority. For middle schools, the peer index was based on the 4th grade reading and math test scores of current students.

that our results represent the impact of group-based teacher performance pay for schools already under accountability pressure.

The timing of program announcement and the selection of schools into the treatment group did not allow much room for behavioral responses to the program in its first year. The school vote took place in November 2007, less than two months before the January reading exam and less than four months before the March math exam.[7] The program continued into the 2008-2009 school year and all but three of the participating schools voted to continue participation.[8] Of the 158 schools that voted to participate in the first year of the program, 87 (55 percent) received bonus payments. The bonus pool averaged approximately $160,500 per school, and totaled $14.0 million in the first year. In the second year of the program, the vast majority of the 151 schools that eligible to receive bonuses earned awards, totaling $27.1 million.

3.3 Data and Empirical Framework

Our analyses focus on schools classified as elementary, middle, and kindergarten through grade 8 (K-8) schools eligible for selection into the bonus program. A total of 181 schools were chosen to participate in the bonus program; 128 schools were placed in the control group. We use publicly available DOE data and measure academic achievement using average math and reading test scores in the 2006-07, 2007-08, and 2008-09 school years (hereafter 2007, 2008, and 2009).

We measure teacher absences, teacher turnover, and the characteristics of newly hired teachers using aggregate statistics from data on individual teachers.[9] In some specifications, we

---

[7] However, even given this short time period, the NYC accountability system led to significant improvements in math and, albeit smaller, improvements in reading (Rockoff and Turner 2010).

[8] Schools that voted no in the first year of the program were not given a second chance to vote on the program. However, we still consider these schools as part of the group originally assigned to the treatment group.

[9] We thank Jonah Rockoff for constructing these aggregate statistics for the purpose of this research.

include information on school demographic characteristics (the percentage of students in each school that are English Language Learners (ELL), special education students, Title I free lunch recipients, and minorities) and each schools performance under the new NYC accountability system, including each school's accountability score and peer index.

3.3.1 Was Randomization Successful?

Our ability to make causal inferences about the effects of teacher incentive pay depends on the success of random assignment. In Table 3.1, we present comparisons of the characteristics of treatment and control groups prior to random assignment, where the treatment group includes schools that were initially selected but did not participate. Treatment and control schools are similar in terms of enrollment, accountability outcomes, student demographics, and teacher characteristics. We find no significant differences between the observable characteristics of treatment and control schools, suggesting a causal interpretation of our results is valid.

We also compare the characteristics of the 309 schools in the experimental sample to other schools in NYC.[10] Given that schools with low peer indices were eligible for selection into the bonus program, it is not surprising that the experimental sample differs from the remainder of NYC schools across a number of dimensions. Schools in the experimental sample had a higher proportion of English Language Learners (ELL), special education, minority students, and students eligible for the Title I free lunch program, as well as lower average math and reading scores. Teachers in the experimental sample had slightly less experience and almost twice as many absences than teachers in other NYC schools. Finally, experimental schools had lower enrollment and fewer teachers than other schools.

[10] We restrict our universe to the 923 schools serving students in kindergarten through eighth grade that received accountability grades and were not charter schools or schools that only serve special education students.

Table 3.2 compares the characteristics of schools by whether or not they voted to participate in the program. Schools voting “no” are largely similar to schools that voted in favor of the program, although, on average, these 25 schools were relatively less disadvantaged and their students had higher test scores.

3.3.2 Regression Framework

We estimate the main effect of the bonus program using the following model:

$$y_{jt} = \delta D_{jt} + \mathbf{X_{jt}}\boldsymbol{\beta} + \varepsilon_{jt} \tag{3.1}$$

where y_{jt} is the outcome of interest for school j in year t, D_{jt} is an indicator selection into the bonus program’s treatment group (regardless of whether the school ultimately participated), $\mathbf{X_{jt}}$ is a vector of school characteristics, and ε_{jt} is an idiosyncratic error term.[11] School observations are weighted by the number of tested students. With successful random assignment, D_{jt} is independent of omitted variables and $\hat{\delta}$ represents the casual impact of the bonus program.

3.4 Results

To preview our estimates of the impact of the bonus program on student achievement, Figures 3.1 and 3.2 display the distribution of average math and reading scores within treatment and control schools in 2007, 2008, and 2009. On average, all NYC schools experienced an increase in average student performance in the two years following the implementation of the program; this pattern holds in the experimental sample. If the bonus program had an impact on test scores, we should observe a rightward shift in the distribution among treatment schools, relative to

[11] Covariates include the outcome measured in 2007, school type indicators (i.e., elementary, middle, or K-8), the percentage of students that are English Language Learners, special education, Title I free lunch recipients, and minorities, and performance under the NYC accountability system (school accountability scores and peer indices).

control schools. The distribution of math and reading scores do not differ significantly between treatment and control schools in either 2008 or 2009.

3.4.1 Math and Reading Achievement

Table 3.3, which displays results from regressions estimating the impact of the program on average math and reading exam scores, confirms these findings. We find little evidence that the program led to increases in math and reading achievement and, if anything, it appears that eligibility to earn bonuses had a negative impact on math achievement. Panels A and B examine the first and second years of the program separately. The point estimates for 2008 are negative and quite small, although precisely estimated.[12] In the second year of the program, eligibility to earn bonuses had no effect on student achievement in reading and a small negative impact on math scores, leading to an approximately 0.08 standard deviation reduction in math achievement.[13]

3.4.2 Group Bonuses and the Free-rider Problem

Teachers should respond to the bonus program by increasing their effort until the expected marginal benefit is equal to the expected marginal cost. However, the probability that a treated school reaches its goal and receives a bonus primarily depends on students' performance on math and reading exams. Thus, the impact of an individual's teacher's effort on her expected bonus is decreasing as the number of teachers with tested students increases.[14] The diffusion of

[12] For instance, our IV estimates reject effects as small as a 0.7 point increase in reading achievement and a 0.2 point increase in math. These effects are quite small in magnitude, given the 2008 student level standard deviation in test scores was 35 points for reading and 31 points for math.

[13] Four schools in the treatment group were closed at the end of the 2008 school year, thus, our sample decreases by four in the second set of regressions. Our 2008 results remain unchanged when we restrict the sample to only include schools open in both 2008 and 2009.

[14] Consider two extremes, a school with only one teacher with tested students and a school with an infinite number of these teachers. In the first case, the teacher will either respond to the program by increasing her effort to the

responsibility for test score gains across many teachers may dilute the incentives of the bonus scheme. Moreover, monitoring may be more difficult in schools with more teachers.

We test for evidence of free-riding by allowing treatment effects on math and reading scores to vary by the number of math and reading teachers, respectively. We only focus on teachers whose students take these exams, rather than the full set of teachers in a school, since only teachers with tested students contribute to the probability that a school earns its bonus.[15] The first set of regressions in Table 3.4 show the main effect of the bonus program on math and reading achievement.[16] We first add an interaction between the number of math/reading teachers (relative to the mean number of such teachers in the sample) and the treatment indicator (columns 2 and 5), and finally, interact treatment status with an indicator for schools in the bottom quartile of the number of teachers with tested students (approximately 10 or fewer teachers in elementary and K-8 schools and 5 or fewer in middle schools). We only present results from specifications that include covariates, however, results are similar when we exclude covariates or instrument for actual treatment with initial assignment.

We find evidence of free-riding. For schools at the bottom of the distribution of the number of teachers with tested students, we estimate a positive effect of the bonus program on math achievement in the first year of the program and a positive, but insignificant effect in the second year, although we cannot reject a test of equality of effects across years. In 2008, the

---

expected level necessary to achieve the school's goal or not respond (if the size of the bonus is less than the cost of exerting this level of effort). In the second case, changes in a given teacher's effort do not affect the probability that the school receives the bonus and it will be optimal for teachers to not respond to the program.

[15] On average, treatment and control group schools have 55 teachers in total, but only 16 teach tested students.

[16] The small number of middle and K-8 schools that are missing information on the number of teachers with tested subjects are excluded.

bonus program resulted in a 3.2 point (0.08 student-level standard deviation) increase in math achievement.[17]

Group-based incentive pay may outperform individual incentives in the case of joint production. If the degree to which teachers work together varies across schools, the bonus program may have been effective in schools with a high level of cooperation between teachers. To proxy for the extent of joint production in a school, we construct a measure of school cohesiveness using teachers' answers to a set of five survey questions prior to the announcement of the bonus program.[18] This measure may also incorporate the degree to which teachers are able to monitor their colleagues. We sum responses across survey questions and standardize the index so it has a mean of zero and standard deviation equal to one. Schools with high levels of cohesion are distinct from those with a small number of teachers with tested students.[19]

Table 3.5 tests for heterogeneity in the impact of the bonus program by school cohesion. We first interact treatment with the linear index (columns 2 and 5) and then interact treatment with an indicator for schools with above average cohesion (columns 3 and 6). The point estimates for schools with below average cohesion are marginally significant and negative in both subjects and both years, while the interaction of treatment and the indicator for above average cohesion is significant, positive, and of greater magnitude. Results suggest that the

[17] Another implication of this finding is that, in schools with a large number of teachers with tested students, the bonus program had a negative impact on student achievement. One explanation is the bonus program crowded out teachers' intrinsic motivation and only in schools where incentives were not diluted by free-riding did the potential monetary rewards lead to increased teacher effort.

[18] These surveys were administered in spring 2007. Questions include: (1) the extent to which teachers report feeling supported by fellow teachers, (2) whether curriculum and instruction is aligned within and across school grades, (3) whether the principal involves teachers in decision making, (4) whether school leaders encourage collaboration, and (5) whether teachers collaborate to improve instruction. We exclude schools with a survey response rate under 10%.

[19] This index has a small, negative, and statistically insignificant correlation with the number of math and reading teachers in a school.

bonus program may have had detrimental effects in schools with low levels of cohesion, and small positive effects on achievement in cohesive schools.

3.4.3 Teacher Effort

A primary motivation for performance-based pay is to provide teachers with incentives to increase effort devoted to raising student achievement. Although we do not directly observe teacher effort, we can measure teacher attendance, which may be correlated with effort decisions and contributes to student achievement (e.g., Miller, Murnane, and Willett 2008; Herrmann and Rockoff forthcoming). We measure teacher absences using aggregate statistics from individual teacher data and estimate models where the dependent variable is the average number of absences taken during the months when schools first learned of their eligibility for the bonus program and when the last exams were taken. If teachers believe that their attendance can affect the probability of bonus receipt by raising student achievement, the program's impacts on absenteeism should be largest over this period.[20] We only examine absences that teachers likely have some control over – those taken for illness and personal reasons.

Table 3.6 presents these results; each column within a panel contains the estimates from separate regressions. The first column examines the effect of the bonus program on absences across all teachers within a school and shows no measurable impact on overall attendance. Column 2 focuses on teachers with tested students, while the third and fourth columns follow the same approach as Table 3.4 and interact the treatment indicator with the number of teachers with tested students (column 3) or an indicator for whether a school falls in the bottom quartile of the number of such teachers (column 4).

---

[20] In the first year of the program, schools learned of their eligibility in November while in the second year, eligibility was known in September. In both years, the last exams occurred in March. Results are robust to alternate definitions of the time period (e.g., November to March in the second year or September to March in the first year).

Program impacts on attendance are not consistent across years. In the program's first year, for schools with a small number of teachers with tested students, attendance increased.[21] Conversely, in the second year of the program, we find positive but insignificant impacts on absenteeism. Finally, we test whether the bonus program had heterogeneous impacts according to initial teacher effort. For instance, initially low effort (high absence) teachers may be the only group with the ability to respond through increasing attendance. Conversely, if *ex ante* high effort teachers believed that achieving the bonus program goals was a high probability event, they may have responded by reducing their effort. However, we find no evidence teacher absenteeism varies along this dimension (available upon request). In the United States, attendance may not be the dimension along which teachers respond to incentive pay.

3.4.4 Bonuses and School Accountability

Finally, we test whether the NYC accountability system, also implemented in 2007, contributed to the bonus program's ineffectiveness. Teachers may have already adjusted their effort or teaching practices in response to the accountability system's incentives. If teachers face decreasing marginal returns or increasing marginal costs to effort, the size of potential bonus payments may not be large enough to induce additional effort.

To evaluate this possibility, we take advantage of the fact that treatment schools face different incentives according to their accountability grades. Both treatment and control schools receiving low grades had additional motivation to improve student test scores, as they faced school closure or principal removal if student achievement did not improve in the following year. Conversely, schools receiving an A on their progress report generally needed to make the smallest gains to receive a bonus, thus, the program may not have provided a large incentive to

---

[21] However, impacts are only significant in schools at the 10^{th} percentile in the distribution of number of teachers (results available upon request).

teachers in treatment schools to alter their behavior. Treatment and control schools in the middle of the grade distribution faced the largest difference in incentives. We test whether treatment effects vary along this dimension, grouping schools into three separate bins by their accountability grades: A, B or C, and D or F. We find no significant differences in treatment effects between these grade groupings or for schools at the center of the grade distribution where the difference in incentives between treatment and control schools are largest (Table 3.7).

3.5 Conclusions

In many sectors, performance-based pay enhances effort, output, and other desirable outcomes. Evidence from Israel and India suggests that properly structured teacher incentive pay programs can benefit students. However, despite substantial expenditures – over $40 million in the program's first two years – the NYC bonus program did not raise student achievement. This paper discusses several features of the NYC bonus program that may have contributed to its ineffectiveness. We provide suggestive evidence that the group-based structure of the program may have been detrimental in the majority of schools where the number of teachers responsible for tested students is large. Conversely, the program improved math achievement in schools with fewer teachers responsible for tested students or a more cohesive group of teachers. A lack of monitoring as well as the diffusion of responsibility for test score gains among many teachers may have diluted the incentives of the opportunity to earn bonuses. Our results are consistent with the long-standing literature in economics on the importance of taking into consideration free-riding, joint production, and monitoring when designing incentive systems and suggest that a one-size-fits-all approach may not be the most effective when implementing incentive pay schemes within a school district.

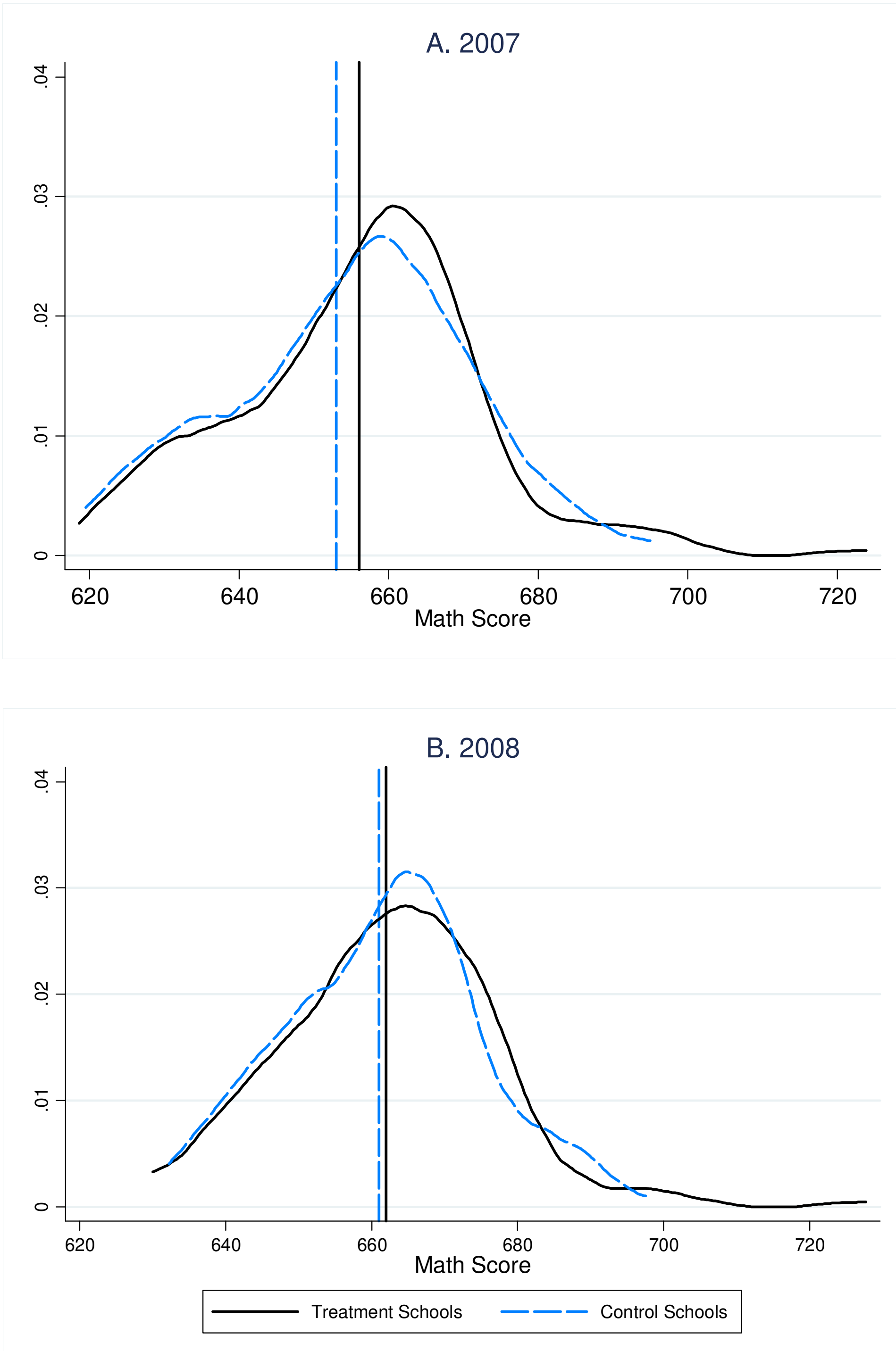
Given that team-based incentives in other contexts resulted in student achievement gains, other features of the NYC program may have also contributed to its ineffectiveness. Neal (2011) suggests that results from economic theory offer valuable insights into optimal incentive design. For instance, an intervention in India utilized a piece-rate payment scheme: teachers or schools received bonus payments for incremental improvements in student achievement (Muralidharan and Sundararaman 2011). This avoids threshold effects of schemes like the NYC bonus program, which dilute incentives for teachers with a probability of bonus receipt approaches zero or one.

Even so, many challenges in designing effective teacher incentive schemes remain. Incentive pay programs that come about as a compromise between school districts and teachers unions' might contain incentives that are so diluted they are destined to fail. Finally, the extensive margin may be most important margin through which teacher pay can improve student achievement. Small-scale teacher incentive pay experiments cannot provide information concerning the general equilibrium effects of overall increase in teacher pay or movement towards performance-based compensation.

Currently, the U.S. government provides significant funding through the Race to the Top program. Eligibility for Race to the Top funding depends on districts' ability and willingness to link student achievement to individual teachers and use this data in teacher evaluations, but grants districts a great deal of discretion in designing performance pay systems. In 2010, 62 school districts and nonprofit groups received over $400 million in funding from the federal Teacher Incentive Fund. Our results underscore the importance of the structure of performance pay in education. Policy innovations in this area should be carefully considered, taking into account personnel economics theory and research.

3.6 Figures and Tables

Figure 3.1: The Distribution of Average Math Scores by Treatment Status



Notes: Vertical lines denote mean math scores for treatment (solid) and control schools (dashed).

Figure 3.1: The Distribution of Average Math Scores by Treatment Status, cont.

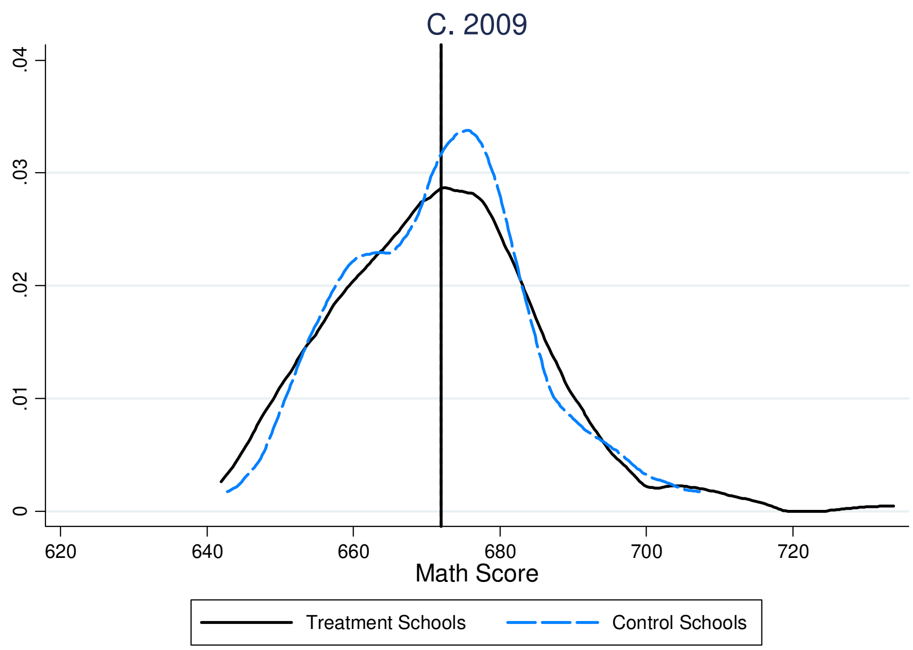
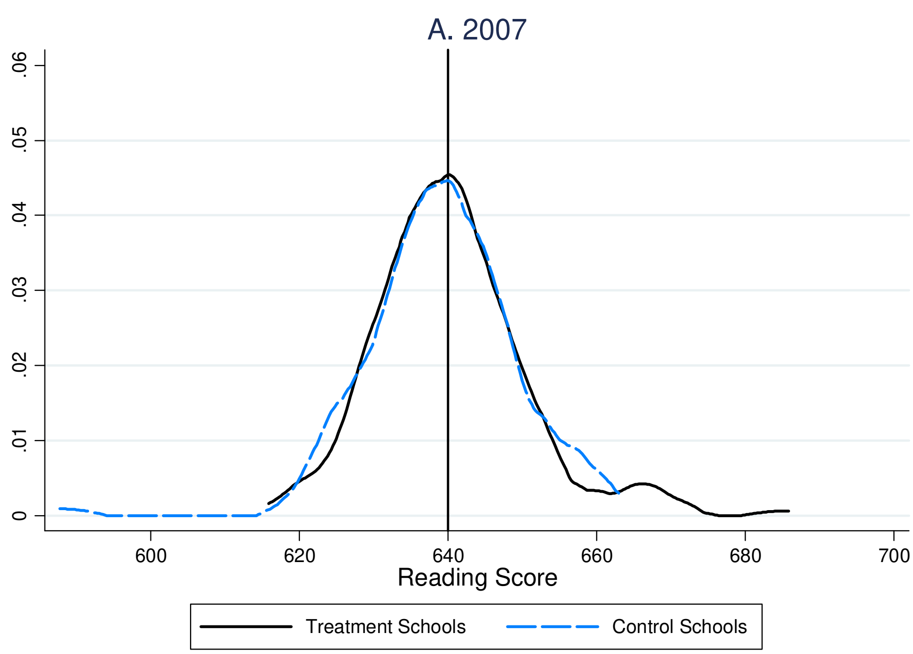
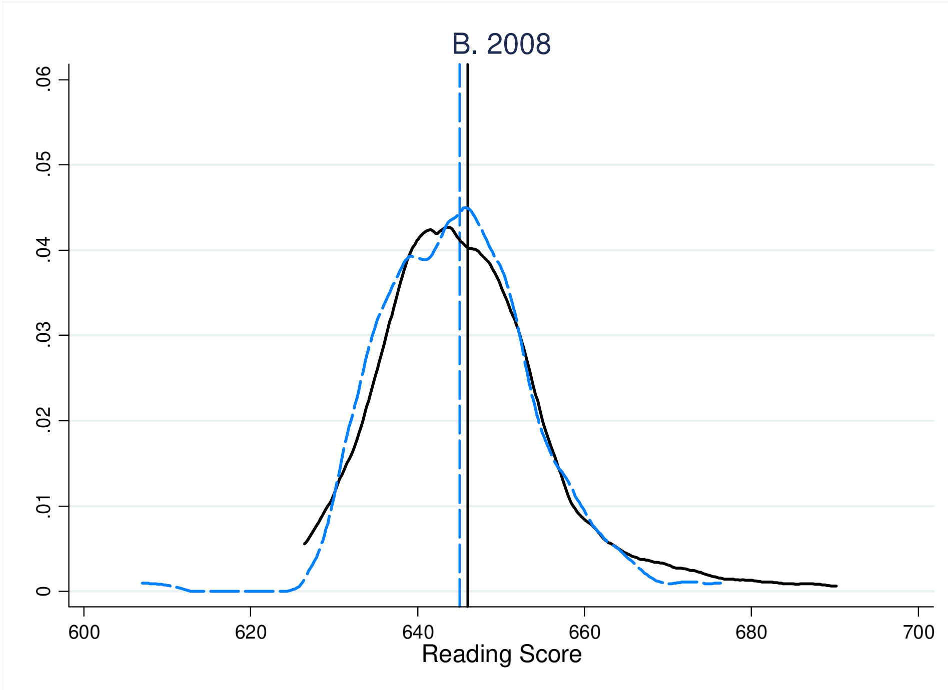


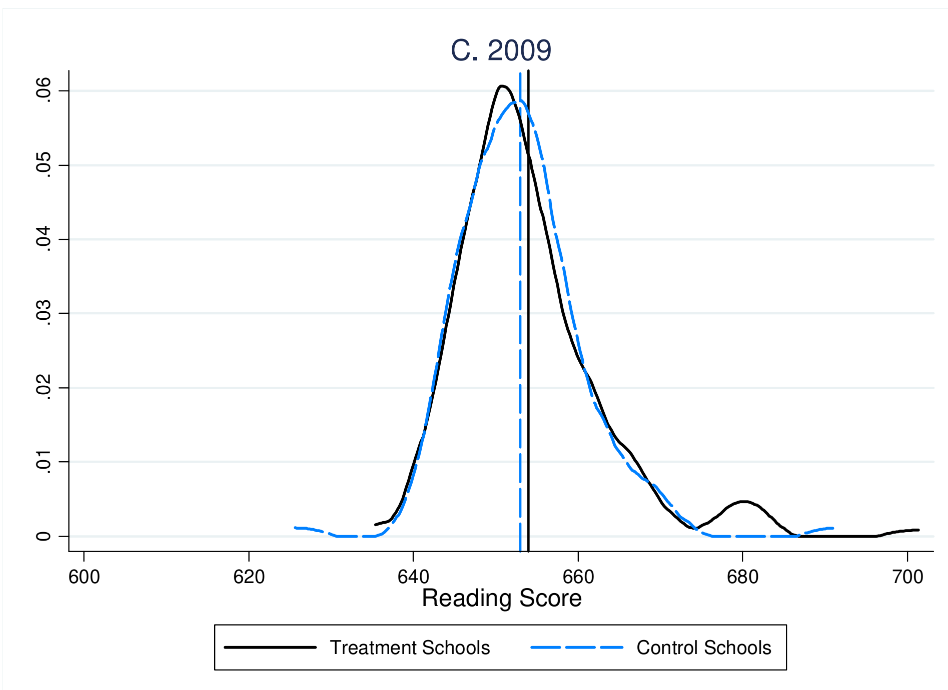
Figure 3.2: The Distribution of Average Reading Scores by Treatment Status



Notes: Vertical lines denote mean math scores for treatment (solid) and control schools (dashed).

Figure 3.2: The Distribution of Average Reading Scores by Treatment Status, cont.





Notes: Vertical lines denote mean math scores for treatment (solid) and control schools (dashed).

Table 3.1: Baseline School Characteristics by Original Assignment to Treatment and Control Groups

	Treatment Schools	Control Schools	Difference	p-value	Non-Experimental Schools
Number of Schools	181	128			614
Average enrollment	558	558	0	0.852	687
Average enrollment, tested grades	363	367	-4	0.912	459
Fraction elementary school	0.62	0.63	-0.01	0.788	0.63
Fraction middle school	0.26	0.27	-0.01	0.586	0.24
Fraction K-8 school	0.12	0.10	0.02	0.452	0.13
School Accountability Outcomes					
Peer index (mean = 0, sd = 1)	-0.91	-0.93	0.02	0.452	0.44
Overall accountability score	52.6	52.1	0.6	0.750	54.6
Target score	66.3	65.9	0.4	0.772	67.8
Student Characteristics					
Average math scale score (2007)	656	655	1	0.497	677
Change in math scale score (2006 to 2007)	10.5	10.3	0.2	0.741	8.7
Average reading scale score (2007)	640	640	1	0.603	660
Change in reading scale score (2006 to 2007)	1.4	1.9	-0.5	0.511	0.9
Fraction English Language Learner	0.19	0.19	0.01	0.614	0.11
Fraction special education	0.12	0.13	-0.01	0.246	0.09
Fraction free lunch	0.87	0.89	-0.02	0.315	0.62
Fraction Hispanic	0.56	0.53	0.03	0.428	0.33
Fraction Black	0.41	0.43	-0.03	0.425	0.29
Fraction White	0.01	0.01	0.00	0.640	0.20
Teacher Characteristics					
Number of teachers	55	55	0	0.952	60
Number of teachers, tested classrooms	16	16	-1	0.431	17
Average years of experience	7.9	8.0	-0.1	0.703	8.6
Average absences/teacher (2007)	7.2	7.0	0.2	0.447	6.7
Average absences/teacher, tested classrooms (2007)	7.4	7.2	0.3	0.377	7.0

Notes: Characteristics measured at beginning of 2007-2008 school year unless otherwise noted; average absences per teacher include absences taken for personal or sick leave.

Table 3.2: Baseline School Characteristics by Participation Vote

	Voted "yes"	Voted "no"	Difference	p-value
Number of Schools	158	25		
Average enrollment	558	574	-16	0.754
Average enrollment, tested grades	364	361	3	0.939
Fraction elementary school	0.61	0.72	-0.11	0.284
Fraction middle school	0.27	0.20	0.07	0.487
Fraction K-8 school	0.13	0.08	0.05	0.508
School Accountability Outcomes				
Peer index (mean = 0, sd = 1)	-0.91	-0.87	-0.05	0.247
Overall accountability score	52.5	55.1	-2.5	0.452
Target score	66.3	68.2	-1.9	0.480
Student Characteristics				
Average math scale score (2007)	655	661	-6	0.102
Change in math scale score (2006 to 2007)	10.7	10.2	0.5	0.704
Average reading scale score (2007)	640	644	-5	0.040
Change in reading scale score (2006 to 2007)	1.7	0.2	1.4	0.316
Fraction English Language Learner	0.20	0.18	0.02	0.549
Fraction special education	0.12	0.12	0.00	0.773
Fraction free lunch	0.88	0.86	0.02	0.608
Fraction Hispanic	0.56	0.54	0.03	0.672
Fraction Black	0.41	0.42	-0.01	0.868
Fraction White	0.01	0.01	0.00	0.772
Teacher Characteristics				
Number of teachers	55	56	-2	0.707
Number of teachers, tested classrooms	15	17	-2	0.237
Average years of experience	7.9	8.4	-0.6	0.163
Average absences (2007)	7.1	7.1	0.0	0.426
Average absences, tested classrooms (2007)	7.0	7.2	-0.2	0.775

Notes: Characteristics measured at beginning of 2007-2008 school year unless otherwise noted; average absences per teacher include absences taken for personal or sick leave.

Table 3.3: The Impact of Teacher Incentives on Student Math and Reading Achievement

	Reading				Math			
	Mean (sd)	(1) OLS	(2) OLS	(3) IV	Mean (sd)	(4) OLS	(5) OLS	(6) IV
A. Year 1: 2007 - 2008								
Treatment	655	-0.876	-0.395	-0.486	672	-1.418	-0.789	-0.970
	(35)	(1.084)	(0.488)	(0.589)	(40)	(1.737)	(0.524)	(0.632)
Observations		309	309	309		309	309	309
B. Year 2: 2008 - 2009								
Treatment	662	-0.852	-0.584	-0.734	680	-1.637	-1.385	-1.740
	(31)	(0.930)	(0.533)	(0.660)	(37)	(1.652)	(0.655)*	(0.813)*
Observations		305	305	305		305	305	305
Additional covariates			X	X			X	X

Notes: + significant at 10%; * significant at 5%; ** significant at 1%; first column displays mean and sd score across all NYC students; each cell denotes a separate regression, dependent variable: school average reading or math score; robust standard errors in parentheses; all regressions weighted by number of students tested in math or reading; additional covariates include: indicators for school level, pre-treatment (2007) math or reading scale score, pre-treatment (2007) peer index and overall accountability score, and pre-treatment (fall 2007) student demographic characteristics: percentage ELL, special education, free lunch recipients, and student race (African American and Hispanic); sample sizes differ across years due to the closure of four schools at the end of the 2007-2008 school year.

Table 3.4: Free-riding and the Impact of Teacher Incentives on Student Math and Reading Achievement

	Reading			Math		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Year 1: 2007-2008						
Treatment	-0.372	0.046	-0.667	-0.871	-0.536	-1.445
	(0.490)	(0.499)	(0.519)	(0.530)	(0.568)	(0.561)*
* Number of teachers *(mean = 0)*		-0.233			-0.176	
		(0.089)**			(0.097)+	
* First quartile of number of teachers			2.044			4.670
			(1.575)			(1.483)**
Treatment effect: schools in first quartile			1.377			3.225
			(1.481)			(1.395)*
Observations	300	300	300	301	301	301
B. Year 2: 2008-2009						
Treatment	-0.579	-0.395	-0.909	-1.297	-0.979	-1.893
	(0.539)	(0.572)	(0.556)	(0.668)+	(0.726)	(0.689)**
* Number of teachers *(mean = 0)*		-0.126			-0.171	
		(0.099)			(0.144)	
* First quartile of number of teachers			2.122			4.826
			(2.067)			(2.579)+
Treatment effect: schools in first quartile			1.213			2.933
			(1.968)			(2.461)
Observations	294	294	294	294	294	294

Notes: + significant at 10%; * significant at 5%; ** significant at 1%; each column denotes a separate regression; robust standard errors in parentheses; measures of the number of reading/math teachers are demeaned; additional controls include: pre-treatment (2007) school test score, school level, peer index, overall accountability score, percentage of students ELL, special education, free lunch recipients, and student race (African American and Hispanic); regressions are weighted by number of tested students; schools with no teachers linked to tested students are dropped; the number of math teachers for schools in the first quartile is less than or equal to: 10 (elementary and K-8 schools), 5 (middle schools); the number of reading teachers for schools in the first quartile is less than or equal to: 10 (elementary and K-8 schools), 6 (middle schools).

Table 3.5: School Cohesion and the Impact of Teacher Incentives on Student Math and Reading Achievement

	Reading			Math		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Year 1: 2007-2008						
Treatment	-0.310	-0.067	-0.912	-0.643	-0.274	-1.288
	(0.492)	(0.511)	(0.592)	(0.529)	(0.549)	(0.674)+
* Cohesion index		0.316			0.797	
		(0.545)			(0.620)	
* Above average cohesion index			1.888			1.987
			(0.968)+			(1.128)+
Treatment effect: schools with above average cohesion			0.976			0.698
			(0.778)			(0.887)
Observations	300	300		301	301	301
B. Year 2: 2008-2009						
Treatment	-0.498	-0.267	-1.153	-1.044	-0.537	-2.326
	(0.533)	(0.554)	(0.661)+	(0.654)	(0.669)	(0.838)**
* Cohesion index		0.406			1.266	
		(0.598)			(0.774)	
* Above average cohesion index			1.982			3.692
			(1.074)+			(1.390)**
Treatment effect: schools with above average cohesion			0.829			1.367
			(0.847)			(1.070)
Observations	299	299	299	300	300	300

Notes: + significant at 10%; * significant at 5%; ** significant at 1%; robust standard errors in parentheses; each column denotes a separate regression; teacher cohesion index mean = 0, sd = 1 across all NYC schools; additional controls include: pre-treatment (2007) school test score, school level, peer index, overall accountability score, percentage of students ELL, special education, free lunch recipients, and student race (African American and Hispanic); regressions are weighted by number of tested students; schools with teacher survey response rate below 10% are dropped.

Table 3.6: The Impact of Teacher Incentives on Teacher Absences Due to Personal and Sick

	All Teachers	Teachers of Tested Students		
	(1)	(2)	(3)	(4)
A. Year 1: 2007-2008				
Treatment	0.001	-0.158	-0.217	-0.156
	(0.091)	(0.146)	(0.148)	(0.163)
* Number of teachers *(mean = 0)*			0.013	
			(0.022)	
* First quartile of number of teachers				-0.236
				(0.390)
Treatment effect: schools in first quartile				-0.391
				(0.352)
Observations	301	301	301	301
B. Year 2: 2008-2009				
Treatment	0.045	0.151	0.203	0.161
	(0.119)	(0.175)	(0.192)	(0.200)
* Number of teachers *(mean = 0)*			0.005	
			(0.032)	
* First quartile of number of teachers				0.158
				(0.621)
Treatment effect: schools in first quartile				0.319
				(0.576)
Observations	294	294	294	294

Notes: + significant at 10%; * significant at 5%; ** significant at 1%; each column within a panel denotes a separate regression; measures of the number of reading/math teachers are demeaned; dependent variable is average absences per teacher taken for personel or sick leave between November and March (Panel A) or September and March (Panel B); additional controls include: pre-treatment (2007) school test score, school level, peer index, overall accountability score, percentage of students ELL, special education, free lunch recipients, and student race (African American and Hispanic); regressions are weighted by number of tested students; schools with no teachers linked to tested students are dropped; the number of teachers for schools in the first quartile is less than or equal to: 10 (elementary schools),11 (middle and K-8 schools).

Table 3.7: Heterogeneity in Impact of Teacher Incentives on Student Math and Reading Achievement by Accountability Grade

	Reading			Math		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	IV	OLS	OLS	IV
A. Year 1: 2007 - 2008						
Treatment*D or F	-3.719	0.175	0.188	-5.292	-0.329	-0.379
	(2.147)+	(1.116)	(1.244)	(3.407)	(1.144)	(1.279)
Treatment*B or C	0.454	-0.733	-0.892	0.944	-0.400	-0.470
	(1.309)	(0.615)	(0.726)	(2.163)	(0.700)	(0.825)
Treatment* A	-1.856	0.063	0.091	-3.703	-1.542	-2.031
	(2.489)	(1.122)	(1.439)	(3.608)	(1.084)	(1.447)
Test A/B = C = D/F (pvalue)	0.231	0.698	0.685	0.238	0.653	0.625
Observations	309	309	309	309	309	309
B. Year 2: 2007 - 2008						
Treatment*D or F	-3.533	-1.924	-3.246	-7.273	-3.326	-5.582
	(2.321)	(1.300)	(2.178)	(4.017)+	(2.388)	(3.924)
Treatment*B or C	-0.488	-0.379	-0.465	-0.686	-0.463	-0.574
	(0.976)	(0.435)	(0.518)	(1.888)	(0.658)	(0.786)
Treatment* A	1.179	-0.091	-0.113	1.511	-0.603	-0.749
	(1.802)	(0.686)	(0.843)	(2.726)	(0.952)	(1.174)
Test A/B = C = D/F (pvalue)	0.277	0.450	0.400	0.193	0.515	0.460
Observations	305	302	302	305	302	302
Additional covariates		X	X		X	X

Notes: + significant at 10%; * significant at 5%; ** significant at 1%; each column within a panel denotes a separate regression; dependent variable: school average reading or math scale score interacted with indicator for school grade; robust standard errors in parentheses; all regressions weighted by number of students tested in math or reading; additional covariates include: prior year scale score, indicators for school level, peer index, overall accountability score, percentage of students ELL, special education, free lunch recipients, and student race (African American and Hispanic); sample sizes differ across years due to the closure of four schools at the end of the 2007-2008 school year and the elimination of an additional three schools that did not receive 2008 accountability grades.

REFERENCES

Angrist, Joshua D. 1993. "The Effect of Veterans Benefits on Education and Earnings." *Industrial and Labor Relations Review* 46(4): 637-52.

Angrist, Joshua D., and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics* 114(2): 533-75.

Bettinger, Eric. 2004. "How Financial Aid Affects Persistence." In *College Choices: The Economics of Where to Go, When to Go, and How to Pay for it*, edited by Caroline M. Hoxby, 207-237. Chicago and London: University of Chicago Press.

Bettinger, Eric P., Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. Forthcoming. "The Role of Simplification and Information in College Decisions: Results from the H&R Block FAFSA Experiment." *Quarterly Journal of Economics*.

Bound, John, and Sarah Turner. 2002. "Going to War and Going to College: Did World War II and the G.I. Bill Increase Educational Attainment for Returning Veterans?" *Journal of Labor Economics* 20(4): 784-815.

Bound, John, Michael F. Lovenheim, and Sarah Turner. 2010. "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources." *American Economic Journal: Applied Economics* 2(3): 129-57.

Bulow, Jeremy I., and Paul Pfleiderer. 1983. "A Note on the Effect of Cost Changes on Prices." *Journal of Political Economy* 91(1): 182-5.

Card, David. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics Volume 3*, edited by Orley C. Ashenfelter and David Card, 1801-63. Amsterdam: Elsevier.

Card, David, David S. Lee, and Zhuan Pei. 2009. "Quasi-Experimental Identification and Estimation in the Regression Kink Design." Princeton University Industrial Relations Section Working Paper 553.

Cellini, Stephanie Riegg. 2010. "Financial Aid and For-Profit Colleges: Does Aid Encourage Entry?" *Journal of Policy Analysis and Management* 29(3): 526-52.

Cellini, Stephanie Riegg, and Latika Chaudhary. 2011. "The Labor Market Returns to a Private Two-Year College Education." Manuscript, George Washington University.

Cellini, Stephanie Riegg, and Claudia Goldin. 2012. "Does Federal Student Aid Raise Tuition? New Evidence on For-Profit Colleges." National Bureau of Economic Research Working Paper 17827.

Charles, Kerwin Kofi, Erik Hurst, and Melvin Stephens Jr. 2008. “Rates for Vehicle Loans: Race and Loan Source.” *American Economic Review* 98(2): 315-20.

Cullen, Julie Berry, and Randall Reback. 2006. “Tinkering Toward Accolades: School Gaming under a Performance Accountability System.” In *Advances in Applied Microeconomics Volume 14: Improving School Accountability*, edited by Timothy J. Gronberg and Dennis W. Jansen, 1-34. Oxford, UK: JAI Press.

Dave, Dhaval M., Nancy E. Reichman, Hope Corman, and Dhiman Das. 2011. “Effects of Welfare Reform on Vocational Education and Training.” *Economics of Education Review* 30(6): 1399-1415.

Dave, Dhaval M., Nancy E. Reichman, and Hope Corman. 2008. “Effects of Welfare Reform on Educational Acquisition of Young Adult Women.” National Bureau of Economic Research Working Paper 14466.

Dee, Thomas S. 2004. “Are there Civic Returns to Education?” *Journal of Public Economics* 88(9-10): 1697-1720.

Deming, David, and Susan Dynarski. 2010. “Into College, Out of Poverty? Policies to Increase the Postsecondary Attainment of the Poor.” In *Targeting Investments in Children: Fighting Poverty When Resources are Limited*, edited by Phillip B. Levine and David J. Zimmerman, 283-302. Chicago and London: University of Chicago Press.

Deming, David J., Claudia Goldin, and Lawrence F. Katz. 2012. “The For-Profit Postsecondary School Sector: Nimble Critters or Agile Predators?” *Journal of Economic Perspectives* 26(1): 139-64.

Dynarski, Susan M. 2003. “Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion.” *American Economic Review* 93(1): 279-88.

Dynarski, Susan M., and Judith E. Scott-Clayton. 2008. “Complexity and Targeting in Federal Student Aid: A Quantitative Analysis.” In *Tax Policy and the Economy, Volume 22*, edited by James M. Poterba, 109-50. Chicago and London: University of Chicago Press.

Fryer, Roland G. 2011. “Teacher Incentives and Student Achievement: Evidence from New York City Public Schools. National Bureau of Economic Research Working Paper 16850.

Fullerton, Don, and Gilbert E. Metcalf. 2002. "Tax incidence." In *Handbook of Public Economics, Volume 4,* edited by Alan J. Auerbach and Martin Feldstein, 1787-1872. Amsterdam: Elsevier.

Gibbons, Robert. 1998. “Incentives in Organizations.” *Journal of Economic Perspectives* 12(4): 115–32.

Goodman, Sarena, and Lesley J. Turner. 2010. "Teacher Incentive Pay and Educational Outcomes: Evidence from the NYC Bonus Program." Program on Education Policy and Governance Working Paper Number 10-07.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klauuw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69(1): 201-9.

Hansmann, Henry B. 1980. "The Role of the Nonprofit Enterprise." *Yale Law Review* 89(5): 835-901.

Hastings, Justine S., and Ebonya L. Washington. 2010. "The First of the Month Effect: Consumer Behavior and Store Responses." *American Economic Journal: Economic Policy* 2(2): 142-62.

Herrmann, Mariesa A., and Jonah E. Rockoff. Forthcoming. "Worker Absence and Productivity: Evidence from Teaching." *Journal of Labor Economics.*

Holmstrom, Bengt. 1982. "Moral Hazard in Teams." *Bell Journal of Economics* 13(2): 324–40.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7: 24–52.

Horn, Laura, and Stephanie Nevill. 2006. Profile of Undergraduates in U.S. Postsecondary Education Institutions 2003-04: With a Special Analysis of Community College Students. NCES 2006-184. U.S. Department of Education, Washington, DC: NCES.

Hotz, V. Joseph, Guido W. Imbens, and Jacob A. Klerman. 2006. "The Long-Term Gains from GAIN: A Re-Analysis of the Impacts of the California GAIN Program." *Journal of Labor Economics* 24(3): 521-66.

Hoxby, Caroline M. 2000. "The Effects of Class Size on Student Achievement: New Evidence from Population Variation." *Quarterly Journal of Economics* 115(4): 1239-85.

Itoh, Hideshi. 1991. "Incentives to Help in Multi-Agent Situations." *Econometrica* 59(3): 611-36.

Jackson, C. Kirabo, and Elias Bruegmann. 2009. "Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers." *American Economic Journal: Applied Economics* 1(4): 1–27.

Jacob, Brian A. 2005. "Accountability, Incentives and Behavior: Evidence from School Reform in Chicago." *Journal of Public Economics* 89(5-6): 761–96.

Jacob, Brian A., and Steven D. Levitt. 2003. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." *Quarterly Journal of Economics* 118(3): 843–77.

Jacobson, Louis, Robert LaLonde, and Daniel G. Sullivan. 2005. "Estimating the Returns to Community College Schooling for Displaced Workers." *Journal of Econometrics* 125(1-2): 271-304.

Jepsen, Christopher, Kenneth Troske, and Paul Coomes. 2009. "The Labor Market Returns to Community College Degrees, Diplomas, and Certificates." University of Kentucky Center for Poverty Research Discussion Paper DP2009-08.

Kane, Thomas J. 1995. "Rising Public College Tuition and College Entry: How Well Do Public Subsidies Promote Access to College?" National Bureau of Economic Research Working Paper 5164.

Kane, Thomas J., and Cecilia Elena Rouse. 1995. "Labor-Market Returns to Two- and Four-Year College." *American Economic Review* 85(3): 600-14.

Langer, Ashley. 2011. "Demographic Preferences and Price Discrimination in New Vehicle Sales." Manuscript, University of Michigan.

Lavy, Victor. 2002. "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement." *The Journal of Political Economy* 110(6): 1286–1317.

Lavy, Victor. 2009. "Performance Pay and Teachers' Effort, Productivity and Grading Ethics." *American Economic Review* 99(5): 1979–2011.

Lazear, Edward P. 2001. "Educational Production." *Quarterly Journal of Economics* 116(3): 777-803.

Lazear, Edward P., and Paul Oyer. Forthcoming. "Personnel Economics." In *Handbook of Organizational Economics*, edited by Robert Gibbons and D. John Roberts. Princeton, NJ: Princeton University Press.

Lee, David S. 2008. "Randomized Experiments from Non-random Selection in U.S. House Elections," *Journal of Econometrics* 142(2): 675-97.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48(2): 281-355.

Li, Judith A. 1999. "Estimating the Effect of Federal Financial Aid on Higher Education: A Study of Pell Grants." Manuscript, Harvard University.

Long, Bridget Terry. 2004. "How do Financial Aid Policies Affect Colleges? The Institutional Impact of the Georgia HOPE Scholarship." *Journal of Human Resources* 39(4): 1045-66.

Lochner, Lance, and Enrico Moretti. 2004. “The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports.” *American Economic Review* 94(1): 155-89.

London, Rebecca A. 2006. “The Role of Postsecondary Education in Welfare Recipients’ Paths to Self-Sufficiency.” *The Journal of Higher Education* 77(3): 472-96.

McCrary, Justin. 2008. “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.” *Journal of Econometrics* 142(2): 698-714.

McPherson, Michael S. and Morton Owen Schapiro. 1991. Keeping College Affordable: Government and Educational Opportunity. Washington, DC: Brookings Institution.

Miller, Raegen T., Richard J. Murnane, and John B. Willett. 2008. “Do Worker Absences Affect Productivity? The Case of Teachers.” *International Labour Review* 147(1): 71–89.

Moretti, Enrico. 2004. “Workers’ Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions.” *American Economic Review* 94(3): 656-90.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. “Teacher Performance Pay: Experimental Evidence from India.” *Journal of Political Economy* 119(1): 39-77.

National Center for Education Statistics. 2011. Digest of Education Statistics, 2010. Washington DC: U.S. Department of Education.

Neal, Derek. 2011. “The Design of Performance Pay in Education.” In *Handbook of Economics of Education, Volume 4*, edited by Eric A. Hanushek, Steve Machin, and Ludger Woessmann, 490-550. Amsterdam: Elsevier.

Neal, Derek, and Diane W. Schanzenbach. 2010. “Left Behind by Design: Proficiency Counts and Test-Based Accountability.” *Review of Economics and Statistics* 92(2): 263-83.

Nielsen, Helena Skyt, Torben Sørensen, and Christopher Taber. 2010. “Estimating the Effect of Student Aid on College Enrollment: Evidence from a Government Grant Policy Reform.” *American Economic Journal: Economic Policy* 2(2): 185-215.

Podgursky, Michael J., and Matthew G. Springer. 2007. “Teacher Performance Pay: A Review.” *Journal of Policy Analysis and Management* 26(4): 909–49.

Riccio, James A., and Daniel Friedlander. 1992. GAIN: Program Strategies, Participation Patterns, and First-Year Impacts in Six Counties. MDRC.

Riccio, James A., Daniel Friedlander, and Stephen Freedman. 1994. GAIN: Benefits, Costs, and Three-Year Impacts of a Welfare-to-Work Program. MDRC.

Rockoff, Jonah, and Lesley J. Turner. 2010. “Short-Run Impacts of Accountability on School Quality.” *American Economic Journal: Economic Policy* 2(4): 119–47.

Rothstein, Jesse. 2008. "The Unintended Consequences of Encouraging Work: Tax Incidence and the EITC." Princeton University Center for Economic Policy Studies Working Paper Number 165.

Seftor, Neil S., and Sarah E. Turner. 2002. "Back to School: Federal Student Aid Policy and Adult College Enrollment." *Journal of Human Resources* 37(2): 336-52.

Singell, Larry D., and Joe A. Stone. 2007. "For Whom the Pell Tolls: The Response of University Tuition to Federal Grants-in-Aid." *Economics of Education Review* 26(3): 285-95.

Springer, Matthew G., and Marcus A. Winters. 2009. "The NYC Teacher Pay-for-Performance Program: Early Evidence from a Randomized Trial." Manhattan Institute Civic Report Number 56.

Turner, Nicholas. 2012. "Who Benefits from Student Aid: The Economic Incidence of Tax-Based Federal Student Aid." *Economics of Education Review*, in press.

Urquiola, Miguel. 2006. "Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia." *Review of Economics and Statistics* 88(1): 171-77.

U.S. Department of Education. 2006. The EFC Formula Guide 2006-2007. Washington, DC.

U.S. Government Accountability Office. 2010. For-Profit Colleges: Undercover Testing Finds Colleges Encouraged Fraud and Engaged in Deceptive and Questionable Marketing Practices (Publication No. GAO-10-948T). Washington, DC.

Weyl, E. Glen, and Michal Fabinger. 2011. "A Restatement of the Theory of Monopoly." Manuscript, University of Chicago.

Yinger, John. 1998. "Evidence on Discrimination in Consumer Markets." *Journal of Economic Perspectives* 12(2): 23-40.