DISSERTATION SERIES 2022:3

Inputs and incentives in education

J. Lucas Tilley

IFAU
INSTITUTE FOR EVALUATION OF LABOUR MARKET AND EDUCATION POLICY

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P O Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

For dissertations, quality is ensured in the traditional way, through the academic review procedure at Uppsala /Stockholm University. This dissertation has not been reviewed by IFAU.

Dissertation presented at Uppsala University publicly examined in Ekonomikum, Hörsal 2, Kyrkogårdsgatan 10, Uppsala, Monday, 7 June 2021 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Olmo Silva (London School of Economics, Department of Geography and Environment and Centre for Economic Performance).

ISSN 1651-4149

Economic Studies 196

J. Lucas Tilley

Inputs and Incentives in Education

Department of Economics, Uppsala University

Visiting address: Kyrkogårdsgatan 10, Uppsala, Sweden
Postal address: Box 513, SE-751 20 Uppsala, Sweden
Telephone: +46 18 471 00 00
Telefax: +46 18 471 14 78
Internet: http://www.nek.uu.se/

ECONOMICS AT UPPSALA UNIVERSITY

The Department of Economics at Uppsala University has a long history. The first chair in Economics in the Nordic countries was instituted at Uppsala University in 1741.

The main focus of research at the department has varied over the years but has typically been oriented towards policy-relevant applied economics, including both theoretical and empirical studies. The currently most active areas of research can be grouped into six categories:

* Labour economics
* Public economics
* Macroeconomics
* Microeconometrics
* Environmental economics
* Housing and urban economics

Additional information about research in progress and published reports is given in our project catalogue. The catalogue can be ordered directly from the Department of Economics.

J. Lucas Tilley

Inputs and Incentives in Education

UPPSALA
UNIVERSITET

Dissertation presented at Uppsala University to be publicly examined in Ekonomikum, Hörsal 2, Kyrkogårdsgatan 10, Uppsala, Monday, 7 June 2021 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Olmo Silva (London School of Economics, Department of Geography and Environment and Centre for Economic Performance).

Abstract
Tilley, J. L. 2021. Inputs and Incentives in Education. *Economic studies* 196. 184 pp. Uppsala: Department of Economics, Uppsala University. ISBN 978-91-506-2874-6.

Essay I. Educational interventions that increase the quality or quantity of school resources may have a limited impact on student achievement if students lack sufficient effort or motivation. A more effective way of raising achievement could be incentivizing students to perform well in school. In this paper, I study whether students respond to non-financial incentives for higher grades, exploiting a reform in Stockholm that made compulsory school grades the sole criteria for admission to high school. Using a difference-in-differences design, I find that the reform increased students' grade point average in compulsory school by 10% of a standard deviation on average. Estimates of the unconditional quantile treatment effects show that the largest shifts occurred just above the middle of the grade distribution, where the performance incentives were strongest. I perform a variety of checks to support the hypothesis that these effects were driven by changes in student effort rather than changes in school grading practices. My findings suggest that behavioral responses from students drive the results. Thus, strengthening the performance incentives implicit in the design of the education system can have a positive effect on student achievement.

Essay II. This paper studies a large-scale educational expansion to evaluate whether shocks to school inputs have an impact on the academic achievement of adult education students. I analyze the spillover effects of a Swedish policy that temporarily doubled enrollment in adult education, thus putting considerable strain on school resources. Because the intervention targeted individuals age 25 and over, my analysis focuses on individuals under age 25 to mitigate concerns that changes in student composition drive my findings. First, I establish that students in regions subject to larger enrollment shocks also experienced stronger negative shocks to school inputs like teacher credentials and per-pupil expenditure. Then, I show that the stronger negative shocks to school inputs coincided with steeper declines in course completion. Taken together, the two sets of results suggest a causal link between school inputs and course dropout.

Essay III. Teachers with stronger academic credentials tend to work in schools with students from more advantaged backgrounds. This paper contributes to an emerging literature on the mechanisms that drive these sorting patterns. With register data covering all college graduates and teachers in Chile between 2007 and 2020, I examine whether earning a more selective teaching degree has a causal effect on the type of schools where graduates teach at the start of their career. For identification, I exploit a college placement mechanism that generates hundreds of admission cutoffs around which access to more selective teaching programs is essentially random. Using the variation around these cutoffs in a regression discontinuity design, I find suggestive evidence that graduating from a more selective teacher program has an effect on teachers' initial job placements. In particular, it increases the probability of working in more urbanized areas and in publicly-subsidized private schools.

Keywords: adult education, human capital, performance incentives, school resources, student achievement, teacher credentials, teacher sorting

J. Lucas Tilley, Department of Economics, Box 513, Uppsala University, SE-75120 Uppsala, Sweden.

ISSN 0283-7668
ISBN 978-91-506-2874-6
urn:nbn:se:uu:diva-439786 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-439786)

Dedicated to my family

Acknowledgments

Despite the periods of self-doubt that naturally come with writing a PhD thesis, part of me isn't at all surprised that I ended up here, studying questions like: Do students respond to incentives, and does access to qualified teachers affect their achievement? After all, I was the kid who felt so inspired by his favorite teachers that he gave lectures to make-believe classrooms full of stuffed animals. The kid who, when his second grade teacher offered an ice cream sandwich to any student who finished five books during Reading Week, spent the entire evening buried in books so that he could get his prize dessert the very next day. Whether or not that kid was bound to end up here, the path to completing this thesis was not always smooth or straight. Many people helped me navigate the ups and downs along the way. Given that my research focuses on factors affecting student achievement, I find it fitting to begin by acknowledging all of you who provided valuable input during my PhD studies.

It goes without saying that I owe my deepest gratitude to my supervisors, Helena Holmlund and Peter Fredriksson. The value that they added to this thesis is immeasurable. Helena—from our very first meeting, you shared and encouraged my enthusiasm for education economics. Over the course of many more, you gave excellent advice about how to improve my research and myself as a researcher. Thank you for your econometric expertise, abundance of patience, and all the extra chats that helped me keep (somewhat) calm and carry on. You went above and beyond what anyone could hope for from their main supervisor. Peter—you, too, were exceptionally generous with your time and extensive knowledge. Even when I found it difficult to explain my problems, let alone solve them, you somehow saw through my confusion and instantly came up with an answer. I also appreciated your meticulous proofreading. If only I had learned to send drafts over earlier so I could have benefitted from it even more!

Before Helena and Peter became my supervisors, Mikael Lindahl and Alex Solís helped steer the course—thank you both for your guidance, especially on what ended up being the final chapter of this thesis. Mikael, I am also grateful that you connected me with David Figlio, who kindly invited me to spend a year at the Institute for Policy Research. I found my time there inspiring.

A number of other researchers provided helpful feedback on early versions of the chapters in this thesis. In particular, I want to thank my opponents, Björn Tyrefors and Alexander Willén, for the excellent discussions during my licentiate and final seminars. My co-evaluator Björn Öckert deserves special mention for his thoroughness and long list of ideas about how to make my papers more convincing. I also want to thank Georg Graetz and Oskar Nordström Skans for their many suggestions over the years, as well as the encouragement they provided when I had presentations.

None of this would have been possible without the Economics Department at Uppsala, where I had the pleasure to work alongside so many remarkable colleagues. First, to the Uppsala Labor Group—I feel fortunate that this group got started while I could still benefit from it. It was great to get feedback at every step of the research process and to see what other PhD students were working on. Second, to the Work Environment Group and the PhD Association—I truly enjoyed collaborating with you to try and make a good workplace even better. I think we succeeded! Third, to the administrative staff—especially Johanna Mörk, whose door was always open when I had an issue—I appreciate all the assistance you provided over the years.

Some of the most important people in the PhD program are those who share the trials and tribulations of the mandatory first-year courses. I lucked out and entered the program with the best possible cohort: Aino-Maija Aalto, Henrik Andersson, Olle Hammar, Dagmar Müller, Maria Olsson, and Paula Roth. The first year was much more bearable because we all recognized the value of collaboration over competition. I appreciated all the laughs we had during lunch (the pronunciation of a certain Finnish word comes to mind!), and of course, all the pancake nights, whether making whipped cream by hand or being served fancy Belgian crepes and waffles. Aino-Maija and Dagmar—I feel especially lucky that we ended up sharing an office during the first year and grew to be such close friends. The PhD program would have been a lot more difficult without your friendship. Thanks for listening when I needed it, but also for the YouTube dances, Grey's nights, and dog pictures when I needed that instead.

So many more of you have made Ekonomikum a better, brighter place. Cristina Bratu—thanks for all the discussions about work, life, work-life balance, and taking it "bird by bird" (even if we haven't quite figured it out yet). You are a brilliant researcher and a good friend. Vivika Halapuu, Olle Hammar, and Melinda Suveg—thanks for being so easy to open up to, both in general and in particular during the past year. You made the job market less daunting and the pandemic less isolating. Jonas Cederlöf, Mohammad Sepahvand, Dmytro Stoyko, and Raoul van Maarseveen—thanks for the amusing conversations and always giving off positive energy (in Mohammad's case, especially when the sun was out!). Gunnar Brandén and Akib Khan—thanks for being awesome office mates, even if for various reasons, we weren't often in at the same time. When we were, I appreciated our chats and the laid-back atmosphere. Ylva Moberg—thanks for being a great first-year mentor, but most of all, for being such a caring friend. I am happy that you (re-)introduced me to Ian Burn and Emma von Essen. Working with the three of you over the past year reminded me just how fun research can be.

Even if they did not contribute directly to my thesis, I definitely got by with a little help from my friends outside of work. A collective thanks to all of you for making sure that I popped out of the academia bubble once in a while. This past year, I was stuck inside another kind of bubble due to the pandemic, and Stina—I am so grateful you chose to be part of mine and offer some much-needed ablenkung. Ellen—you deserve an award for enduring all my nonsense. I hope you enjoy our mischief as much as I do. Johan—it feels emptier in Uppsala without you, but visiting you and Dasha always boosts my spirits and gives me the right perspective on things. Robert—I am happy to have found someone as passionate (crazy, even?) about running as I am. I enjoyed all the weekends in Hågadalen and Hammarskog. Adam, Amelia, Antonia (my BFFIB!), Greg, Hilary, Kara, and Ri—I see you nowhere near enough. When I do, it feels like nothing has changed.

Last but the opposite of least, my family has provided the most important input throughout my life. That remains true to this day, even with an ocean (and for now, a bunch of travel restrictions) between us. I want to thank my "bonus family" for putting up with another quick-to-debate Tilley. Charlie and Livy—for all the good talks and good food. Deborah—for the support and assurance someone was there to love Sadie. Ron—for the recipes, memes, and tolerating all the video calls with my Mom. Mom—somehow distance brought us even closer together. Thank you for the countless hours of conversation and for always believing in my abilities, even when I didn't. Dad, or "Chief"—I am called the Mini-Chief for a reason, and it is not just my height issues! You understand my peculiarities better than most. Thank you for the numerous ways you have supported me over the years. Much of that support enabled me to get here. Brad and Sanna—I appreciate that you let me take over your guest room for extended periods of time whenever I came back to the states. My best memories these past few years are from our visits. Finally, Kennedy and Emelyn—thanks for giving me a true break from research and keeping my mind occupied with everything from the names of the Paw Patrol characters to deep toddler wisdom like "It matters because it's not important." You have taught me what *is* important.

Uppsala, April 2021

Contents

Introduction

The importance of education is by now indisputable. Early research on the economics of education documented robust positive correlations between years of schooling and outcomes like income and employment (see, e.g., Becker, 1983; Rosen, 1977; Willis, 1986). These studies offered suggestive evidence that education improves people's economic well-being, but they could not rule out the possibility that people with higher earnings potential simply decide to stay in school longer.[1] By the 1990s, a so-called "credibility revolution" in empirical microeconomics led to the use of more sophisticated methods and better research designs, and with them, it was possible to establish that education in fact has a causal effect on labor market outcomes (see, e.g., Angrist and Krueger, 1991; Angrist and Pischke, 2010; Card, 1999). More recently, the scope of these studies has broadened to consider the possible non-monetary benefits of education as well. We now know that education not only raises earnings and employability; it also has a positive impact on a wide range of non-economic outcomes including health, family stability, civic engagement, and crime (see, e.g., Dee, 2004; Heckman et al., 2018; Lochner and Moretti, 2004; Oreopoulos and Salvanes, 2011).

While the benefits of education are well established, finding effective ways to raise educational attainment and achievement has proven to be a much more contentious matter. Two questions have been especially divisive: first, do school inputs like class size and per-pupil expenditure have any effect on student outcomes (see, e.g., Hedges et al., 2016)? Second, even if these factors do have an effect, is investing a significant amount of money in educational interventions like class-size reductions worth it, particularly in the case of older or more disadvantaged students? It has been difficult to reach a consensus both on the type of policies that are effective at raising achievement levels in general, but also on the type of policies that are effective at reducing achievement gaps across socioeconomic groups in particular (see, e.g., Heckman et al., 2003).

[1] For readers who have not sat through an Econometrics 101 class, the usual lesson goes something like this: Correlation does not imply causation! Individuals who complete more years of schooling probably differ in other important ways that improve their employment outcomes—-for example, their level of motivation or ambition.

Researchers on one side of the debate point to the evidence that the most critical period of skill formation occurs early in life and that significant differences in cognitive and non-cognitive abilities emerge even before individuals reach schooling age (see, e.g., Cunha and Heckman, 2007). They argue that educational investments made at older ages have significantly lower rates of return than those made at younger ages, particularly for disadvantaged students (see, e.g., Heckman, 2006). Thus, more emphasis should be placed on family policies early in life, especially if the goal is closing achievement gaps. As far as education interventions, they suggest that policies promoting school competition and choice are more likely to improve school quality than policies targeting school inputs (see, e.g., Heckman et al., 2003). This argument was seemingly in line with a large body of research showing that school-based inputs like class size, teacher qualifications, and per-pupil spending do not have a strong or consistent relationship with student performance, at least once family characteristics are taken account into (see, e.g., Hanushek, 1997).

On the other side of the debate, researchers have maintained that the reviews of the literature on school inputs have been misleading (see, e.g., Card and Krueger, 1998; Hedges et al., 2016; Jackson, 2020) and that the most credible evidence suggests that interventions like class-size reductions can have large effects, especially for vulnerable students (see, e.g., Krueger and Whitmore, 2001). Indeed, a growing number of studies suggest that school resources do matter and that later-age interventions can have meaningful impacts on student outcomes (see, e.g., Carrell and Sacerdote, 2017; Jackson et al., 2016; Rea and Burton, 2020a,b).

Against this backdrop, my dissertation provides evidence on how various features of the education system affect the outcomes of students in the later stages of their education. It consists of three independent chapters focusing on individuals in secondary and post-secondary schooling. The first two are set in the Swedish context and explore the effect of policy interventions on academic achievement. More specifically, the first chapter studies whether strengthening the performance incentives built into high school admission policies has an effect on students' grades, while the second assesses whether policy-induced shocks to school resources affect the performance of adult education students. In the final chapter, my analysis shifts to the Chilean context and studies whether attending more prestigious teacher education programs has an effect on the type of schools where teachers find their first jobs. All three chapters rely on rich administrative data and use quasi-experimental methods that allow me to identify causal effects. In the remainder of this introduction, I provide a short overview of the topics addressed in each thesis chapter and briefly present my main findings.

The main topics and findings of this thesis

I. Performance incentives in high school admission policies

How much a student learns in school depends in part on the amount of effort that they exert. Their academic performance can suffer if they become disengaged and invest inadequate time or effort in their schoolwork. This may happen because they lack accurate information about the value of their education and the ways in which their hard work could benefit them in the future (see, e.g., Fryer, 2016). Alternatively, they may understand that performing well in school will ultimately pay off, but when deciding how to allocate their time and effort—say, watching a hockey game with friends versus studying hard and trying to ace their next maths test—they may place too little weight on the long-run benefits of studying and instead prefer the short-term enjoyment of doing something fun (see, e.g., Bettinger and Slonim, 2007).

In such instances, providing students with incentives to perform well in school may boost their level of motivation, effort, and hopefully, their achievement and learning. The idea that individuals respond to incentives is a fundamental concept in economics, and when it comes to getting kids to work harder in school, it almost seems like common sense that it works: parents have long known that if you want a kid to do something (like their homework), it helps to offer them something that they want in return (like iPad privileges)! However, incentives can have a downside; if learning becomes all about external rewards, students may lose their own intrinsic motivation to learn, and their performance may actually be harmed in the long run (see, e.g., Gneezy et al., 2011).

A number of studies have investigated whether students respond to incentives to work harder or perform better in school, and if so, whether their responses persist once the incentive is removed. Most of these studies are experimental in nature and look at the effect of monetary incentives (see, e.g., Angrist et al., 2009; Fryer, 2011; Leuven et al., 2010). The researchers randomly assign students to some treatment—for example, paying them a certain amount of money for each assignment they submit, for each high grade they earn, or for passing important achievement tests–and then measure how they respond relative to a group of students who were not provided with the same incentives. The evidence on the effectiveness of these interventions is mixed, but some suggest that—if properly designed—there is scope for financial incentives to boost student effort and achievement (for an overview, see Gneezy et al., 2011). However, providing these financial incentives is a very costly intervention. There is thus increasing interest in understanding whether students might respond to non-financial incentives, which are cheaper[2] and often viewed as less controversial than policies that pay students to learn.

The first chapter of my thesis contributes to a small but growing literature studying whether students respond to non-financial incentives to perform well in school. It is one of only a few studies to consider how students respond to the performance incentives embedded in the design of school admission systems. I evaluate a reform in Stockholm that made compulsory school grades the sole criteria for admission to high school, thereby strengthening the incentives for students to earn high grades so that they could get into their desired school. I find that the reform had a positive effect on students' compulsory school grades, with the largest shifts occurring just above the middle of the grade distribution, where the performance incentives were strongest. Compared to previous studies, I provide a more rigorous investigation of the mechanisms behind the results. In particular, I perform several checks to separate between student effort and grade inflation. All of the tests support the hypothesis that the increase in grades is driven by increased student effort rather than increased grading leniency. However, it is unclear that these effects persist and translate into long-term achievement gains. The reform raised students' probability of graduating from high school, but conditional on graduation, there was no effect on high school grades.

II. School resources in adult education

Even if students have sufficient motivation and work hard in school, their achievement gains may be limited if they do not have adequate access to resources that aid their learning. Thus, the second chapter of my thesis addresses a classic question in education economics: whether school resources matter for student outcomes. While this question has been studied extensively at lower levels of education, I provide the first causal evidence for adult education students. This is a particularly relevant population to study in light of recent studies showing that adult education can improve the labor market outcomes of participants (see, e.g., Blundell et al., 2020) and the fact that policymakers have begun to embrace "lifelong learning" as a key policy tool to cope with technological changes on the labor market. In addition to focusing on an understudied population, another key contribution of my paper is that it highlights how large-scale interventions intended to increase access to education or improve the quality of schooling can actually have harmful side effects that dampen the potential benefits of the intervention.

[2] There may still be direct administrative costs or indirect costs in terms of unintended side effects.

In the paper, I study a Swedish policy that suddenly doubled enrollment in the adult education sector, creating an additional 100,000 spots in municipal adult education over a period of just two years. The expansion was so rapid and massive that it placed considerable strain on school resources. The central government provided generous subsidies to help municipalities cover the increased financial costs associated with the reform, but the extra money was stretched thin in areas that experienced larger expansions. Moreover, the reform created such great demand for new teachers that municipalities had no choice but to hire uncertified, inexperienced teachers, resulting in significant drops in the average qualification of the teaching staff. My analysis considers how these shocks to school inputs affected the academic achievement of adult learners.

One potential complication when studying an educational expansion is that the composition of students changes as a result of the reform. In particular, the average ability of students is likely to decline with the influx of new students. Thus, any observed changes in academic achievement might be a consequence of changes in average ability rather than changes in school inputs. To deal with this concern, I exploit the fact that the expansion I study primarily targeted individuals age 25 to 55 and restrict my analysis to individuals age 24 and under.

I study the relationship between school inputs and student outcomes by performing two complementary analyses. First, I leverage geographical variation in the intensity of adult education expansion and show that, on average, regions experiencing greater enrollment shocks also faced stronger negative shocks to school resources such as per-pupil expenditure and teacher credentials. Then, I show that the stronger negative resource shocks coincided with steeper increases in course dropout among adult learners age 24 and under. When studying the dynamics of the effects over time, the pattern of changes in school inputs and student outcomes is quite similar. Taken together, the two findings are highly suggestive of a causal link between school inputs and course dropout.

III. Prestige of teacher education and teacher sorting patterns

A central tenet of many education systems is that schooling can and should be the "great equalizer" of different life circumstances, providing all students the same chance to succeed regardless of their background. Teachers are widely regarded as one of the most important school resources for ensuring students' success (see, e.g., Rivkin et al., 2005). Although it is difficult to identify the exact attributes that make some teachers more effective than others, researchers have consistently found that good teachers can have sizable impacts on student learning and can help even the most disadvantaged pupils overcome achievement deficits due to poor learning environments at home. Moreover, the positive effects of having a good teacher seem to persist into adulthood across dimensions such as college attendance and earnings (see, e.g., Chetty et al., 2014). Given this evidence, policymakers have made it an explicit goal to guarantee that all schools, irrespective of the students they serve, have a qualified teaching staff.

Despite this goal, a large number of U.S.-based studies show that qualified teachers are unevenly distributed across schools in ways that disadvantage the most vulnerable students (see, e.g., Goldhaber et al., 2015). In the third chapter of my thesis, I document that this same sorting pattern exists in Chile. Using rich administrative data covering all college graduates and teachers, I show that relative to graduates of less prestigious programs, graduates of more prestigious teaching programs are less likely to work in schools that serve students from disadvantaged socioeconomic backgrounds, including public schools and rural schools. Then, I study a potential mechanism underlying this sorting pattern: that earning a more selective degree could have a causal impact on the type of schools where teachers work at the start of their career. This might be the case if, for example, schools use teachers' academic credentials as a signal of their ability when deciding who to hire (see, e.g., Altonji and Pierret, 2001).

In order to isolate the causal effect of graduating from a more selective teaching program, I exploit features of the centralized admissions process at Chile's most prestigious universities. The admissions process generates hundreds of admissions cutoffs around which access to more selective programs is essentially random. My empirical strategy compares the early labor market outcomes of applicants who just pass the admission threshold to their preferred teaching program to the outcomes of applicants who just miss the admission threshold and instead end up in a less selective teaching program. I study the characteristics of the schools where teachers find their first teaching job, focusing on several school attributes across which teachers tend to be unevenly distributed: urbanicity; public versus private ownership; and socioeconomic composition of the student body. My results show that attending and graduating from a more selective teaching program does have a causal impact on teachers' initial job placements. In particular, it increases the probability of working in more urbanized areas and in publicly-subsidized private schools. The findings of this chapter suggest that schools value teachers' academic credentials in the hiring process. Future research could explore whether this is the case because attending a more selective teaching program actually makes an individual a more effective teacher.

References

Altonji, J. G. and Pierret, C. R. (2001). Employer Learning and Statistical Discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.

Angrist, J. and Krueger, A. (1991). Does Compulsory School Attendance Affect Schooling and Earnings? *The Quarterly Journal of Economics*, 106(4):979–1014.

Angrist, J., Lang, D., and Oreopoulos, P. (2009). Incentives and Services for College Achievement: Evidence from a Randomized Trial. *American Economic Journal: Applied Economics*, 1(1):136–163.

Angrist, J. D. and Pischke, J.-S. (2010). The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics. *Journal of Economic Perspectives*, 24(2):3–30.

Becker, G. S. (1983). *Human capital: a theoretical and empirical analysis, with special reference to education*. Midway Reprint. The Univ. of Chicago Pr, Chicago, 2. ed ; repr edition. OCLC: 249879014.

Bettinger, E. and Slonim, R. (2007). Patience among children. *Journal of Public Economics*, 91(1):343–363.

Blundell, R. W., Salvanes, K. G., and Bennett, P. (2020). A Second Chance? Labor Market Returns to Adult Education Using School Reforms. *SSRN Electronic Journal*.

Card, D. (1999). The Causal Effect of Education on Earnings. In *Handbook of Labor Economics*, volume 3, pages 1801–1863. Elsevier.

Card, D. and Krueger, A. B. (1998). School Resources and Student Outcomes. *The Annals of the American Academy of Political and Social Science*, 559:39–53.

Carrell, S. and Sacerdote, B. (2017). Why Do College-Going Interventions Work? *American Economic Journal: Applied Economics*, 9(3):124–151.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, 104(9):2633–2679.

Cunha, F. and Heckman, J. (2007). The Technology of Skill Formation. *The American Economic Review*, 97(2):31–47.

Dee, T. S. (2004). Are there civic returns to education? *Journal of Public Economics*, 88(9):1697–1720.

Fryer, R. G. (2016). Information, non-financial incentives, and student achievement: Evidence from a text messaging experiment. *Journal of Public Economics*, 144:109–121.

Fryer, Jr., R. G. (2011). Financial Incentives and Student Achievement: Evidence from Randomized Trials. *The Quarterly Journal of Economics*, 126(4):1755–1798.

Gneezy, U., Meier, S., and Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, 25(4):191–210.

Goldhaber, D., Lavery, L., and Theobald, R. (2015). Uneven Playing Field? Assessing the Teacher Quality Gap Between Advantaged and Disadvantaged Students. *Educational Researcher*, 44(5):293–307. Publisher: American Educational Research Association.

Hanushek, E. A. (1997). Assessing the Effects of School Resources on Student Performance: An Update. *Educational Evaluation and Policy Analysis*, 19(2):141–164.

Heckman, J. and Mosso, S. (2014). The Economics of Human Development and Social Mobility. *Annual Review of Economics*, 6(1):689–733.

Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, 312:1900–1902.

Heckman, J. J., Humphries, J. E., and Veramendi, G. (2018). Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking. *Journal of Political Economy*, 126(S1):S197–S242.

Heckman, J.J., Krueger, Alan B., and Friedman, Benjamin M. *Inequality in America : What Role for Human Capital Policies?*, volume 2003. MIT Press. Publisher: The MIT Press.

Hedges, L. V., Pigott, T. D., Polanin, J. R., Ryan, A. M., Tocci, C., and Williams, R. T. (2016). The Question of School Resources and Student Achievement: A History and Reconsideration. *Review of Research in Education*, 40:143–168.

Jackson, C. K. (2020). Does School Spending Matter? The New Literature on an Old Question. In *Confronting Inequality: How Policies and Practices Shape Children's Opportunities*, pages 165–186.

Jackson, C. K., Johnson, R. C., and Persico, C. (2016). The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. *The Quarterly Journal of Economics*, 131(1):157–218.

Krueger, A. B. and Whitmore, D. M. (2001). The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR. *The Economic Journal*, 111(468):1–28. Publisher: [Royal Economic Society, Wiley].

Leuven, E., Oosterbeek, H., and van der Klaauw, B. (2010). The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment. *Journal of the European Economic Association*, 8(6):1243–1265.

Lochner, L. and Moretti, E. (2004). The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports. *The American Economic Review*, 94(1):65.

Oreopoulos, P. and Salvanes, K. G. (2011). Priceless: The Nonpecuniary Benefits of Schooling. *Journal of Economic Perspectives*, 25(1):159–184.

Rea, D. and Burton, T. (2020a). Clarifying the Nature of the Heckman Curve. *Journal of Economic Surveys*. https://doi.org/10.1111/joes.12359.

Rea, D. and Burton, T. (2020b). New Evidence on the Heckman Curve. *Journal of Economic Surveys*, 34(2):241–262.

Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73(2):417–458.

Rosen, S. (1977). Human capital: A survey of empirical research. In *Research in Labor Economics*, pages 3–40. JAI Press, Greenwich.

Willis, R. J. (1986). Chapter 10. Wage determinants: A survey and reinterpretation of human capital earnings functions. In *Handbook of Labor Economics*, volume 1, pages 525–602. Elsevier.

I. The effect of higher-stakes grades on student achievement

Acknowledgments: I would like to thank Helena Holmlund and Peter Fredriksson for their guidance on this paper. I am also thankful for helpful feedback from Alexander Willén, Björn Öckert, Raoul van Maarseveen, Cristina Bratu, Jonathan Stråle, Yoko Okuyama, and members of the Uppsala Labor Group.

1. Introduction

Students' performance in compulsory school has important consequences for their later-life outcomes. Achievement measures like course grades and test scores influence how long an individual stays in school, their field of study, and ultimately, their career path and earnings. Thus, it is essential to understand which policy tools can boost student achievement. After decades of inconclusive evidence on the effectiveness of school inputs like class size and per-pupil spending,[1] policymakers are increasingly interested in the use of performance incentives.

In this paper, I evaluate whether strengthening the performance incentives built into school admission policies has an effect on student achievement. Prior research on the use of student-based incentives has remained somewhat limited. Most of the literature evaluates experimental interventions that manipulate students' financial incentives to perform well in school.[2] Only a few papers study how students respond to the incentives embedded in the design of the education system, and whether changing these incentives has unintended side effects.

A high school admission reform in Stockholm, where the admission process is fairly competitive and positions at elite inner-city schools are highly sought after, provides an ideal setting to investigate whether students react to incentives for higher grades. Prior to the reform, the admission process distinguished between admission to a specific school and admission to a specific track.[3] Earning high grades increased students' likelihood of being admitted to their preferred track, but not necessarily their preferred school; rather, assignments to oversubscribed schools were made on the basis of commuting distance. After the reform, students applied directly to a particular school-track combination, and admission decisions were based exclusively on prior grades. The new rules thus strengthened the incentive for students to perform well in school, as earning better grades now increased their chance of attending their preferred high school.

To evaluate the impact of the reform on students' grades, I use a difference-in-differences strategy that exploits geographical variation in the implementation of the reform. Stockholm municipality, which switched to grades-based admissions in year 2000, serves as the treated group, and the neighboring municipalities in Stockholm county, which did not implement the change, serve as the comparison group. My main estimates indicate that, on average, students'

[1] See, e.g., Hedges et al. (2016) and Jackson (2020) for an overview.
[2] Gneezy et al. (2011) provides a good review of the literature.
[3] Tracks can be academic (e.g., the social sciences program or the natural sciences program) or vocational (e.g., the construction program or the handicraft program).

grade point average (GPA) increased around 10% of a standard deviation as a result of the reform. I also examine the effects across the GPA distribution using the unconditional quantile regression method described in Firpo et al. (2009). The results reveal that there were sizable treatment effects across the entire distribution, but the effect size was not uniform: the estimates roughly follow an inverted U-shape, peaking at 15% of a standard deviation just above the median. This pattern indicates that the largest shifts occurred in the middle of the grade distribution, where most admission cutoffs are and thus the performance incentives are strongest.

An important question from a policy perspective is whether these achievement gains actually reflect increased effort and learning, or whether they simply reflect a change in grading practices. The admission reform not only strengthened students' incentives to work harder and perform better in school; given the high level of school competition in Sweden, the reform also strengthened the incentives for schools to inflate grades in an attempt to market themselves to prospective students. Thus, I perform several analyses to investigate the mechanisms behind my results, with a focus on disentangling the student effort channel from the grade inflation channel.

I provide two direct checks on the grade inflation channel using data from standardized exams. These checks exploit the fact that teachers are required to take students' standardized exam grades into account when setting final course grades in math, English, and Swedish. First, with a bunching estimator similar to the one in Chetty et al. (2011), I measure the amount of manipulation around grade cutoffs on national exams in math.[4] Although the grading criteria are centrally determined, the exams are graded by students' teachers. Thus, I can gauge the prevalence of grade inflation by comparing the level of test score manipulation in the treated and control regions. If grade inflation drives my results, we would expect to see a higher degree of test score manipulation in the treated region. However, I find no such evidence. As a second check for grade inflation, I study the effect of the admission reform on teachers' grading leniency using an "inflation index" that measures students' course grades relative to their exam grades. In line with the results of the bunching analysis, I find no increases in grading leniency as a result of the reform.

I perform two additional analyses to corroborate my findings on the mechanisms. First, I show that the effects of the reform do not differ depending on the degree of school competition faced by a student's compulsory school, even though schools subject to more competition should have larger incentives to

[4] I do not perform the bunching analysis for English and Swedish because there are only letter grades for these exams, and the bunching analysis exploits grade cutoffs in the distribution of numerical test scores.

inflate grades. Second, I show that the treatment effects are strongest among the group of students who have a higher likelihood of applying to an academic track, and who thus have the largest potential benefit of earning higher grades. All of the results render support for the hypothesis that changes in student effort drive the increases in grades after the admission reform.

Another important question is whether the short-run effects on compulsory school grades persist or fade out over time. To address this question, I use my main difference-in-differences model to estimate the impact of the reform on student achievement in high school. In line with an earlier study by Söderström (2006), who analyzed only one year of post-reform data, I find that the admission reform had no impact on grades in mandatory high school courses. However, this result should be interpreted with caution, given that high school grades are missing for any student who fails to graduate. When I estimate the effect of reform on the likelihood of graduation, I find that it increased the probability of earning a high school diploma by about 2.4 percentage points, or 3.4% relative to baseline. This suggests that there could be some longer-lasting effect on student learning.

The findings of this study contribute to several strands of literature. First, they relate to the literature on school interventions aimed at boosting student achievement. Most of this research focuses on the effects of school inputs like class size and school spending (see, e.g., Angrist et al., 2019; Fredriksson et al., 2013; Jackson et al., 2016). The studies have mixed findings, and it is often unclear whether the benefits of the interventions exceed their costs. Thus, policymakers are seeking alternative policy measures to improve student achievement. I contribute by studying an intervention that targets a different kind of input in the education production function—namely, the amount of effort that students exert. Several empirical studies support the standard assumption that student effort matters for achievement (Metcalfe et al., 2019; Stinebrickner and Stinebrickner, 2008), but how to increase student effort remains an open question.

A growing literature addresses this question by studying the effectiveness of incentives in educational settings. The majority of these studies evaluate experimental interventions that offer financial rewards to students who exert more effort or perform better in school (see, e.g., Angrist et al., 2009; Bettinger, 2011; Levitt et al., 2016). Only a few look at non-financial rewards, and even fewer do so in a non-experimental context (see, e.g., Baumert and Demmrich, 2001; Grove and Wasserman, 2006; Jalava et al., 2015). The results of these studies vary, and it is unclear whether they can be generalized to higher-stakes settings. I contribute to the literature by studying a more salient change in the incentives built into the education system. Several other studies examine grades-based incentives (Hvidman and Sievertsen, 2019; Koerselman, 2013) and school admission policies in particular (Fajnzylber et al., 2019; Haraldsvik, 2014; Molin, 2019). However, I go beyond these studies by providing a more rigorous investigation of the mechanisms driving the results, which is crucial for understanding the policy implications. For example, I provide direct evidence on teachers' grading practices by measuring the extent to which they manipulate test scores.

The rest of this paper proceeds as follows. The next section provides background information on the Swedish education system and the Stockholm admission reform. Section 3 discusses the empirical framework and model, followed by a description of the data in section 4. Section 5 presents the results of my linear and non-linear difference-in-differences models, as well as some robustness checks to support my findings. Section 6 and 7 provide checks on the mechanisms. Section 8 presents evidence on longer-run effects, and section 9 concludes.

2. Institutional background

I begin by providing key institutional details about the Swedish education system, with a focus on compulsory school grades and the high school admission process. Then, I give an overview of the Stockholm admission reform that I use to study how students respond to incentives for higher grades. All of the facts that I present apply specifically to my period of study (1994–2004), and some have changed in later years.[5]

2.1 Compulsory school grades in Sweden

Compulsory education in Sweden follows a national curriculum and lasts through grade nine, when students are 16 years old. At most grade levels, teachers evaluate student performance with development talks and written assessments rather than course grades. Students receive course grades just four times: once at the end of each term in eighth and ninth grade (Skolverket, 2005). The first three assessments have low stakes, given that the grade point average (GPA) reported on a student's school-leaving certificate (*slutbetyg*) only includes the grades awarded in the spring semester of grade nine. The final assessment has higher stakes depending on where a student lives and what high school program they wish to attend (see section 2.2).

[5]For example, students now receive grades for the first time at an earlier age, and these grades are assigned on a different scale.

Teachers can use discretion when setting final grades, but national criteria guide their grading practices. The criteria changed substantively several years prior to the admission reform that I study. Through the 1996/97 school year, teachers were supposed to assess each student's performance relative to the performance of other students. The grade scale moved in integer increments from one to five, and grades in each subject were supposed to approximate a normal distribution with an average of three.[6] By contrast, from the 1997/98 school year onward, teachers were supposed to assess each student's performance on a criterion-referenced scale where each grade reflected the student's ability to meet specific learning goals. The new scale had only three grades—Pass (*G*); Pass with Distinction (*VG*); and Pass with High Distinction (*MVG*)—and the National Agency for Education (*Skolverket*) outlined the specific learning objectives that students had to achieve in order to receive a certain grade.[7] In my empirical analysis, I deal with the change in grading scale by standardizing each student's grade point average with respect to the national distribution for their graduation cohort. I also show that my results are unchanged if I restrict my sample to students who graduated from compulsory school once the new grading criteria were in place.[8]

By the end of the time period that I study, teachers were supposed to set final course grades in math, English, and Swedish with the aid of standardized tests in each subject. Between 1998 and 2002, the three subject tests were piloted on grade nine students in a sample of municipalities,[9] and in 2003, it became mandatory for all schools to administer the tests at the end of ninth grade. The National Agency for Education provides a manual with detailed grading criteria, but the tests are not centrally graded. Rather, teachers grade their own students' exams on the same G-VG-MVG scale used for course grades. Teachers are supposed to adhere to the national guidelines when setting the exam grade; moreover, they are supposed to take a student's exam grade into account when setting their final course grade. However, the rules are not binding, and evidence suggests that some teachers are too lenient and

[6]The official requirement was that the grades two and four should be more common than the grades one and five. The approximation to a normal distribution was supposed to hold at the national level, though some teachers assumed it should apply at the local level (Skolverket, 2005).

[7]Through the 1999/2000 school year, the official guidelines only outlined the criteria for grades of Pass and Pass with Distinction. From the 2000/01 school year onward, the National Agency for Education also provided criteria for the highest grade, Pass with High Distinction.

[8]I do not make this restriction in my main analysis, because doing so would leave me with only two pre-reform years to investigate parallel trends.

[9]The content of the test and the grading criteria were still under development and therefore changed during the trial period. The test was revised by pedagogical experts based on feedback from teachers.

inflate grades (Diamond and Persson, 2016; Skolverket, 2007). In section 6, I evaluate the extent to which this grading leniency varies between regions and over time to investigate the mechanisms driving my results.

2.2 High school admissions in Sweden

Upon successful completion of compulsory school, students are entitled to three years of tuition-free high school education.[10] High school is organized into different tracks with either an academic or vocational orientation. There are four academic tracks (e.g., the social sciences program) and thirteen vocational tracks (e.g., the construction program). Students have the right to apply to any track in their municipality of residence, and they can only apply outside this region under special circumstances—for example, if their home municipality does not offer their desired field of study, or if their home municipality has an admission agreement with another region.

The admission rules for high school vary by school type, by region, and over time. Admission to independent high schools, which are privately run but publicly funded, grew increasingly common in the late 1990s and early 2000s. By national law, admission to tracks at independent schools has always been made exclusively on the basis of compulsory school grade point average. However, a large majority of students attend the public school system, in which admission decisions are governed by both national and municipality-specific regulations, and the importance of compulsory school grades can vary.

In the past, the public school system drew an important distinction between admission to a track and admission to a school. The admission process worked as follows: First, students submitted an application to the admission center in their municipality of residence. On the application, they listed up to six tracks in order of preference. If the track was offered at more than one public school in the municipality, students could also list their preference for a particular school. Next, the admission center ranked all applicants to a particular track in order of their compulsory school grade point average. If the number of applicants exceeded the number of available slots, national regulations required that the admission center accept students to the track strictly on the basis of compulsory school grade point average.[11] After admitting students to tracks, the admission center had to determine which school the student was going to attend. Whenever possible, municipalities admitted students to the most-

[10] Since the 1997/98 school year, a student must receive a passing grade in math, English, and Swedish to successfully complete compulsory school.

[11] Gymnasieförordning, 1992:394, 6 kap. Behörighet, urval och förfarandet vid antagning [High school ordinance, 1992:394, chapter 6. Eligibility, selection and the admission procedure].

preferred school listed on their application. However, in the case of oversubscribed programs, some selection criteria had to be applied. Each municipality was free to determine their own rules. Some municipalities placed students in the school closest to their home, while others decided which school a student would attend using a lottery or compulsory school grades.

2.3 Overview of the Stockholm admission reform

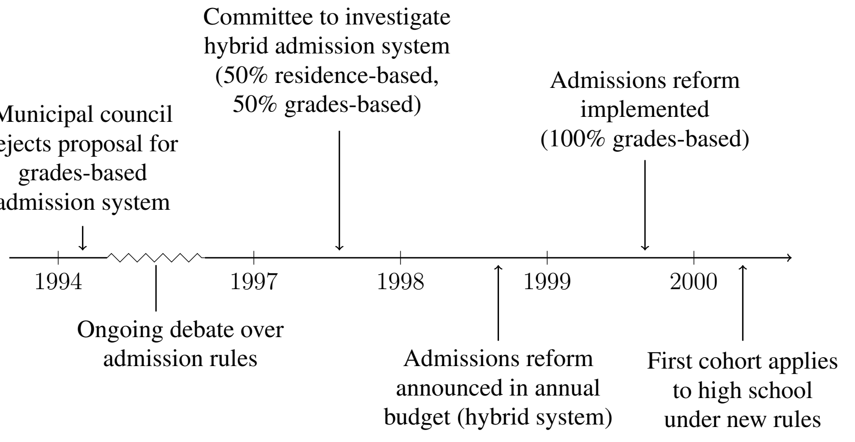
As discussed above, municipalities have some control over the admission rules for their public high schools. In particular, they can decide which school placement mechanism to use if multiple public schools offer the same high school track and the number of applicants exceeds the number of available spots.

Local politicians in Stockholm municipality began debating the merits of different school placement mechanisms in the mid-1990s (see Figure 1 for a timeline). At the time, Stockholm used a proximity-based rule that assigned students to schools according to their commuting distance. Students who lived closer to a school had priority over students who lived farther away. Opponents of the system argued that it reinforced residential segregation and disadvantaged high-achieving students who lived outside the city center, where the most popular schools were located. As an alternative, they proposed that public school admissions become strictly grades-based.

The municipal council rejected the first several policy proposals submitted by the opposition. However, in November 1997, both sides unanimously supported a motion that commissioned an investigative committee to review a possible reform of the admission rules. The proposed plan was a hybrid system with 50% grades-based and 50% proximity-based admissions. According to a newspaper report, the new admission system was expected to be implemented for the 1999 admission round pending the committee's findings (Orrenius, 1997).

The debate over high school admission rules intensified and became a prominent public issue in 1998, which was an election year. The right-wing coalition campaigned on a promise to implement grades-based admissions if elected, while the left-wing coalition began to back-peddle and reaffirm their earlier support for proximity-based school assignment. In October 1998, the right-wing coalition won the election and immediately wrote the high school admission reform into their annual budget. They initially planned to implement the hybrid system that was under evaluation by the investigative committee. However, they ultimately agreed on a broader reform that fully eliminated

Figure 1. Timeline of the admission reform in Stockholm municipality.



Notes: This figure provides a timeline of the high school admission reform in Stockholm municipality. There was an election in October 1998, which the pro-reform coalition won.

proximity-based admissions in favor of grades-based admissions.[12] The reform was officially implemented in October 1999 and became effective immediately. Students who started ninth grade in fall 1999 and graduated from compulsory school in spring 2000 became the first cohort to apply to high school under the new rules.

Previous research has shown that, contrary to its goal, the admission reform actually increased segregation by immigration background, and as expected, by compulsory school grades (Söderström and Uusitalo, 2010). There is also evidence that the reform had positive long-run effects on students' expected income and type of college education (Molin, 2019), despite the absence of a short-run effect on high school grades (Söderström, 2006). However, to my knowledge, there is no evidence on whether the stronger performance incentives created by the reform had an effect on earlier-age achievement. My analysis addresses that question.

[12] An important part of the reform was that high schools were supposed to develop a particular profile, i.e., they had to specialize in a small number of tracks. The pro-reform coalition argued that the more profiled a high school was, the less that proximity-based admissions made sense. This was the motivation for the full switch to grades-based admissions.

3. Empirical framework and method

The aim of this paper is to evaluate whether performance incentives in the education system have an impact on student achievement. I provide evidence from a high school admission reform that replaced residence-based school assignment with grades-based admissions. Before proceeding to the empirical setup, I elaborate on how the reform strengthened the incentives for higher grades. Then, I present the linear and non-linear difference-in-differences models that I use to estimate the average effect of the reform and to study the effects across the grade distribution.

3.1 Changes in incentives due to the admission reform

Prior to the Stockholm admission reform, compulsory school grades determined admission to a particular high school track, but not admission to a specific high school. For any track with more applicants than available spots at the municipal level, students had to earn sufficiently high grades to gain admittance. However, admittance to a specific track did not guarantee admittance to a specific school.[13] If a student wished to attend their preferred track at one of the more popular schools, they had to live close to the school to be guaranteed a spot. This was true no matter how high their compulsory school grades were, given that school placement was proximity-based rather than grades-based. Thus, if students lived too far away from their desired school, they had nothing else to gain by earning higher grades as long as their GPA was high enough to get into their desired track. In other words, the proximity-based admission system created no incentive for students to perform better in school once their grades were above the expected cutoff for the track that they wanted to attend.

After the admission reform, students applied to a specific school-track combination, and compulsory school grades simultaneously determined admission to both the track and the school. Under the reasonable assumption that students or their parents place value on the school they attend—i.e., that going to their preferred school is a "reward" they care about—the new rules strengthened the incentives to earn higher grades.[14] For students who lived far away from their preferred school, earning sufficiently high grades could now guarantee them access to the school. At the same time, students who lived close to their preferred school could no longer take for granted that they would be admit-

[13] An exception is when the track was offered at only one school: in that case, admission to a track implies admission to the school.

[14] See e.g., Black (1999), Figlio and Lucas (2004), and Burgess et al. (2019) for evidence that parents value school quality and school choice.

ted just by crossing the GPA threshold for their preferred track. Under the new system, they had to compete with *all* applicants, not just those who lived nearby; thus, unless they improved their grades, they could be crowded out by higher-achieving students who lived far away.

In addition to strengthening the incentives for students to earn high grades, the admission reform also strengthened the incentives for schools and teachers to award higher grades. The Swedish school system in general, and a large municipality like Stockholm in particular, is characterized by a high degree of school competition. Students have the right to attend any compulsory school within their home municipality, and the municipality must pay each school a per-student voucher for each student that attends. In an attempt to attract more students and the money that comes with them, schools might put pressure on their teachers to inflate grades, given that ninth-grade GPA is the primary way that compulsory schools can market themselves to the parents of prospective students.[15] While these competitive pressures and thus the incentives to inflate grades already existed in the pre-reform period, the reform strengthened them by raising the marketing value of grades.

A key contribution of my paper is that I provide a thorough investigation of the mechanisms behind my results, with a focus on whether the observed effects are driven by changes in grading practices rather than changes in student performance. I postpone discussion of the methods that I use to disentangle the mechanisms until sections 6 and 7, when I also explain the intuition behind the analyses and present my findings.

3.2 Difference-in-differences models

I identify the effect of the admission reform on compulsory school grades using the following difference-in-differences model:

$$\begin{aligned} Grade9GPA_{i,y} = \beta_0 + \beta_1 Treat_i + \beta_2 Post_y + \lambda Treat_i \times Post_y + \gamma_0 Anticipate_{99} \\ + \gamma_1 Treat_i \times Anticipate_{99} + \mathbf{X'}_i \psi + \varepsilon_{i,y} \end{aligned} \tag{1}$$

where the index y denotes the calendar year that student i graduates from ninth grade. The outcome variable, $Grade9GPA_{i,y}$, equals a student's standardized

[15] Given the evidence that parents care about what school their child attends (see previous footnote), the parents themselves could also put pressure on the teachers to inflate grades.

GPA in the national distribution for their graduation cohort.[16] The anticipation indicator, $Anticipate_{99}$, equals one for the cohort that graduated in spring 1999. The post-reform indicator, $Post_y$, equals one for cohorts that graduated under the new high school admission rules, i.e., from year 2000 onward. The treatment group indicator, $Treat_i$, equals one for individuals in Stockholm municipality and zero for individuals in the surrounding municipalities in Stockholm county. The vector $\mathbf{X'}_i$ is a set of individual controls. Finally, $\varepsilon_{i,y}$ is the error term.

There are two important points regarding the model specification. First, I assign each student i to the treated or control group based on their pre-reform municipality of residence. This method of treatment assignment allows me to deal with concerns that high-achieving students in the control regions may re-locate to Stockholm municipality in order to take advantage of the admission reform. Accounting for strategic re-locations is important to avoid bias in my estimates, but I also show that my results hold if I assign students to treatment based on their municipality of residence at the start of ninth grade or the municipality where their compulsory school is located.

Second, I "dummy out" the last pre-reform cohort using the $Anticipate_{99}$ indicator in order to allow for an anticipation effect. It is reasonable to expect an anticipation effect given the timeline of the reform (see Figure 1). There was a tentative plan to change the admission rules for this cohort as far back as November 1997. Although the plans were delayed and would not become effective in time for their graduation, the reform was officially written into the municipal budget shortly after they started ninth grade.

The difference-in-differences model in equation (1) compares the evolution of grades in Stockholm municipality to the evolution of grades in the neighboring municipalities. The main identifying assumption is that, in the absence of the reform, compulsory school grades would have evolved similarly in the two regions. Under this assumption, the parameter λ captures the average effect of the reform.[17] If students react by increasing their effort or teachers react by inflating grades, we would expect the estimate of λ to be positive and significantly different from zero.[18]

In addition to estimating the average effect of the reform, I study how the change in admission rules shifted different parts of the grade distribution. I use

[16] The results are robust to other ways of normalizing the outcome, such as percentile ranking by graduation cohort or standardizing with respect to the GPA distribution in the non-treated regions. These results are available upon request.

[17] It can be considered an intent-to-treat effect, because I assign students to the treated and control groups based on their municipality of residence prior to the reform.

[18] On the other hand, it is possible that the reform decreases students' intrinsic motivation or increases students' stress and performance anxiety. In this case, λ might be negative.

the unconditional quantile regression method described in Firpo et al. (2009) and adapted to the difference-in-differences setting in Havnes and Mogstad (2015). For each percentile of the pre-reform distribution in the treated region, I define an indicator variable equal to one if an individual's standardized GPA exceeds the standardized GPA at that percentile. Then, I use the following linear probability model to estimate how the admission reform shifts the cumulative distribution function:

$$\mathbb{1}[Grade9GPA_i > Grade9GPA_\tau] = \alpha_0 + \alpha_1 Treat_i + \alpha_2 Post_y + \delta Treat_i \times Post_y \\ + \theta_0 Anticipate_{99} + \theta_1 Treat_i \times Anticipate_{99} + \mathbf{X'}_i \psi + \varepsilon_{i,y} \tag{2}$$

where $Grade9GPA_\tau$ denotes the GPA at the τ-th percentile of the pre-reform grade distribution in the treated region. Finally, I re-scale the estimate of δ with an estimate of the unconditional density at $Grade9GPA_\tau$. That is, I calculate the unconditional quantile treatment effect at each percentile τ as follows:

$$\frac{\hat{\delta}}{\hat{f}(Grade9GPA_\tau)} \tag{3}$$

where I estimate the unconditional density $f(Grade9GPA_\tau)$ using a kernel density estimator.

4. Data and sample description

The main data source is the grade nine registry from Statistics Sweden. The database covers all students who finish ninth grade in a given year and includes information on their final grade point average (GPA), as well as their final course grades in each subject. I define my outcome of interest as an individual's standardized GPA within the national grade distribution for their graduation cohort.

I merge the database of compulsory school graduates to the multi-generation registry. This registry provides information on key demographic variables that correlate with student achievement, including students' gender, age, immigration background, birth order, and number of siblings. Importantly, the multi-generation registry also contains personal identifiers for the students' parents. With these identifiers, I can link parents to the Integrated Database for Labor Market Research (*LOUISE*). From the LOUISE database, I obtain information on other relevant demographic variables such as parents' level of education, field of study, employment status, and labor income. I measure the variables at the end of the calendar year prior to the child's graduation, i.e., during the autumn term that the child attends grade nine.

The LOUISE database provides another piece of information that is essential for my analysis: a student's pre-reform municipality. Recall that I define an individual's treatment status according to their pre-reform municipality rather than the municipality where they graduate from compulsory school or reside in ninth grade, so that I do not have to be concerned about strategic re-locations in response to the reform. For all individuals, I define the pre-reform municipality as the municipality where their mother lived in 1993. I use the mother's residence as a proxy for the student's residence because the vast majority of my sample is too young to appear in the population registry that year.[19]

Having merged the data, I construct my estimation sample as follows: First, I restrict the sample to all students who graduate compulsory school between 1994 and 2004. Second, I restrict the sample to students who were residents of Stockholm county at the end of 1993. Finally, I drop a negligible share of students—around 1.5%—for whom key demographic variables are missing. The final sample consists of 178,116 individuals. Table A.1 in the appendix shows the descriptive statistics for the sample by treated and control group. I also present the same descriptive statistics for the full population of compulsory school students in Sweden between 1994 and 2004 for comparison. The treated and control groups are relatively similar on most characteristics, with the most notable difference being the fact that more students in the treated region have unmarried parents. Their parents are also slightly more educated on average.

The last row of Table A.1 shows that, averaged over the entire sample period, students in the treated group perform better in compulsory school than students in the control group: their average GPA is about 13% of a standard deviation higher (0.178 compared to 0.051). To preview the results of my difference-in-differences analysis, Table 1 on the following page goes a step further: for each region, it reports the average GPA in the pre- and post-reform period, as well as the descriptive statistics at other key points of the grade distribution. In the control region, there are few changes between the two periods, whereas in the treated region, standardized GPA increases at every percentile, with a particularly large shift at the median. Figure 2 provides a more detailed graphical illustration. Comparing the pre- to post-reform shifts observed in the two panels, the differences are most striking in the middle to upper end

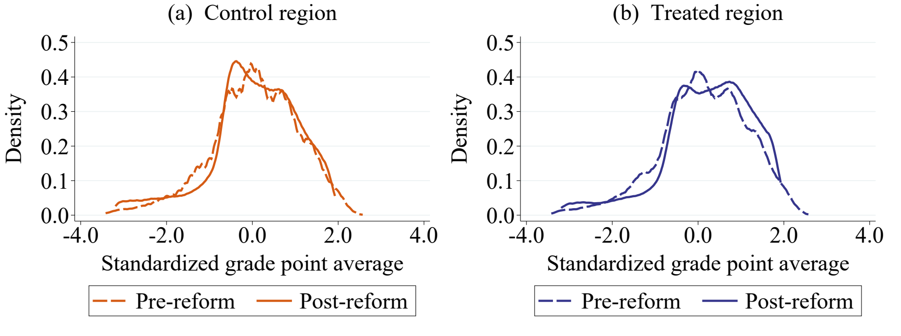
[19] Individuals first appear in the registry the year that they turn 16.

Table 1. Distribution of grade point average pre- and post-reform.

	Pre-reform		Post-reform	
	Control	Treated	Control	Treated
10th percentile	-1.287	-1.176	-1.247	-1.015
25th percentile	-0.559	-0.469	-0.491	-0.357
Mean	0.055	0.132	0.046	0.225
Median	0.109	0.136	0.111	0.329
75th percentile	0.783	0.852	0.777	0.996
90th percentile	1.329	1.426	1.318	1.461
Observations	63,678	29,877	55,897	28,664

Notes: The treated group refers to students who lived in Stockholm municipality at the start of the study period, and the control group refers to students who lived elsewhere in Stockholm county. The pre-reform period lasts from 1994 through 1999, and the post-reform period lasts from 2000 through 2004.

Figure 2. Distribution of grade point average pre- and post-reform.



Notes: This figure shows how the distribution of compulsory school grades changed from the pre-reform to post-reform period in the control region (left panel) and treated region (right panel). The dashed lines show the distribution of grades for ninth graders who graduated prior to the admission reform (1994-1999), and the solid lines show the distribution of grades for ninth graders who graduated after the admission reform (2000-2004).

of the distribution. Panel (a) shows that in the control region, the post-reform distribution (solid line) essentially lies on top of the pre-reform distribution (dashed line) from the midpoint on. However, in the treated region, there is considerably more density at the upper end of the distribution after the reform, suggesting large increases in student grades between the two periods.

5. The effect of the admission reform on compulsory school grades

The descriptive statistics presented in the previous section suggest that compulsory school grades increased when high school admissions became exclusively grades-based. In this section, I formally investigate the effect of the reform by estimating the difference-in-differences model in equation (1). Table 2 presents the main results. Each column reports the estimate of the parameter λ from slightly different model specifications in which I introduce covariates, fixed effects, and time trends to control for factors that could affect the evolution of grades. If the estimate of λ is positive, this indicates that compulsory school grades increase when grades have higher stakes in the high school admission process.

The first row of Table 2 shows that the change in admission rules indeed had a positive effect on students' grades. In column 1, I start with a baseline specification that does not include any control variables. The baseline estimate implies that compulsory school GPA increased 11% of a standard deviation for students in the treated region compared to the control region following the implementation of the admission reform. The effect is quite stable across the different model specifications. Between columns 1 and 2, the point estimate drops slightly to 10% of a standard deviation when I add individual-level control variables to the model. Once I have controlled for these individual characteristics, the estimates are essentially unchanged when I add municipality fixed effects to control for time invariant differences across regions (column 3); time fixed effects to capture shocks or trends that impact treated and control regions equally (column 4); and municipality-specific time trends in grades (column 5). Because the fixed effects and time trends do not affect the main point estimate or improve precision, I exclude them from the models that I run in the remainder of the paper, unless otherwise specified.

As a point of comparison, it helps to relate my estimates to the impact of other educational interventions such as class size reductions. The findings in the literature have been mixed, but a U.S.-based study by Krueger and Whitmore (2001) is often used as a benchmark. They found that reducing class

Table 2. Effect of the admission reform on compulsory school grade point average.

	(1)	(2)	(3)	(4)	(5)
Treat × Post	0.111	0.100	0.099	0.098	0.095
	(0.018)***	(0.015)***	(0.015)***	(0.015)***	(0.016)***
Wild bootstrap p-value	0.039	0.034	0.041	0.041	0.058
Pre-reform mean in treated region	0.132	0.132	0.132	0.132	0.132
Observations	178,116	178,116	178,116	178,116	178,116
Individual controls	No	Yes	Yes	Yes	Yes
Pre-reform municipality fixed effects	No	No	Yes	Yes	Yes
Graduation year fixed effects	No	No	No	Yes	Yes
Municipality-specific linear trend	No	No	No	No	Yes

Notes: The first entry in each column reports the point estimate of λ from equation (1) when including different control variables and time trends in the model. Individual controls include gender, age, foreign background, parents' marital status, number of siblings, birth order, parents' income, and parents' years of education. In column 5, I control for municipality-specific trends using a linear projection of the pre-reform trend. Standard errors are clustered at the municipal level and shown in parentheses. I denote the significance level of the estimate with stars: *** for p-value < 0.01; ** for $p < 0.05$; and * for $p < 0.10$. Because there are only 25 clusters, I also report the p-value when I perform inference using the wild-cluster bootstrap procedure suggested by Cameron et al. (2008).

size by seven students improved student achievement by 0.22 standard deviations, which is in line with the results found by Fredriksson et al. (2013) in the Swedish context.

5.1 Identifying assumptions and robustness checks

Parallel trends assumption

My difference-in-differences model identifies the average causal effect of the reform under the assumption that average grades would have evolved similarly in the treated and control regions if the admission rules had not changed. To check the credibility of this assumption, I examine the evolution of grades over time in the two groups. If the assumption holds, student outcomes should move parallel to one another in the years leading up to the reform and diverge thereafter.

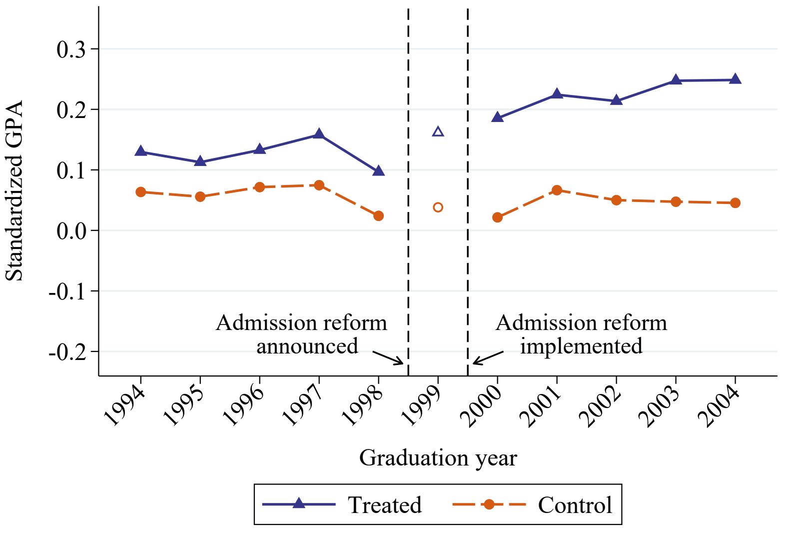
The results in Figure 3 show that this is the case. The solid blue line in panel (a) plots the average GPA by year for the treatment group, and the dashed orange line plots the corresponding averages for the control group. The parallel trends assumption can be assessed by visual inspection of panel (a), but I also provide a more formal test for differential pre-trends in panel (b). This panel reports the yearly estimates from an event study specification in which I replace the post-reform indicator in equation (1) with a set of time dummies. The yearly coefficients are normalized with respect to 1998, when the reform was announced. They trace out the relative trends in the pre-treatment period and also show how the effect of the reform varies over time in the post-reform period.

In panel (a), we see that prior to the announcement of the reform, compulsory school grades evolved in a completely parallel fashion in the two groups. Panel (b) confirms this, showing that all point estimates leading up to the announcement of the reform are statistically indistinguishable from zero. After the announcement of the reform, there is a jump in grades in the treated group, consistent with an anticipation effect. Once the reform is implemented, there is another jump, and grades continue to increase relative to the control group.

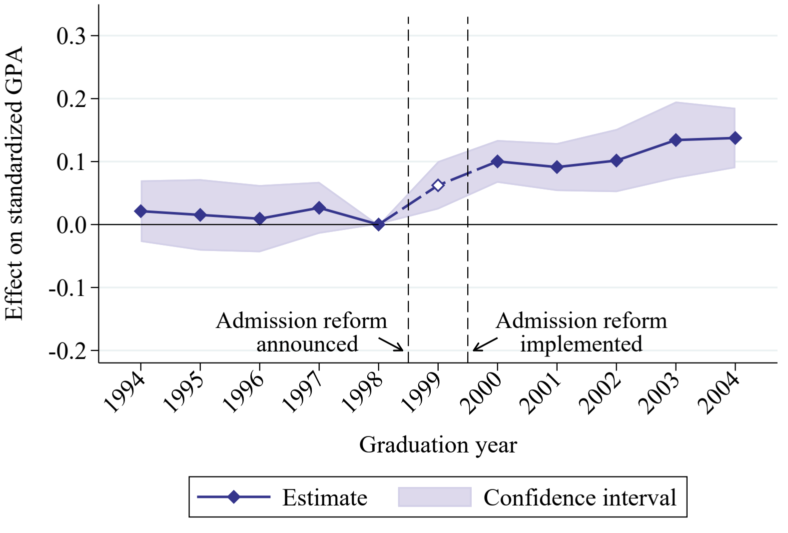
A crucial question is whether it is reasonable to expect such immediate effects after the announcement of the reform. If teachers are inflating grades, then they can react immediately, and they have an incentive to do so as soon as the reform is announced. However, given the institutional setup, students can also have an immediate response. Recall that only the final course grades from ninth grade count in the grade point average used for high school admission; their performance in earlier grades does not directly weigh in. Thus, simply

Figure 3. Evolution of standardized grade point average over time.

(a) Descriptive plot



(b) Event study estimates



Notes: Panel (a) plots the mean standardized grade point average (GPA) for each cohort of ninth graders from spring 1994 to spring 2004. Panel (b) plots the corresponding estimates from an event study specification of equation (1) in which I replace the post-reform indicator with a set of time dummies. I omit the time dummy for 1998 so that all annual coefficients are measured relative to the year prior to the announcement of the admission reform. The first vertical line marks the date that the reform was announced in the municipal budget, and the second vertical line marks the date that the reform went into effect.

by adjusting their effort in ninth grade, they can have a big impact on their grades.[20]

Group composition changes

Another concern is that group composition may have changed over the study period in a way that is related to student achievement. A clear violation would be if motivated students from the control group re-located to the treatment group to take advantage of the new admission rules. To deal with this concern, I have assigned all individuals to the treated and control groups based on their pre-reform municipality of residence. Nevertheless, it is still possible that changes in group composition drive the results. While I cannot rule out the possibility that group composition changed in some unobserved way, I test for changes in observed variables to lend credibility to this assumption. In Table A.2, I report the relevant regression coefficients, standard errors, and p-values from estimation of my main difference-in-differences specification when using the listed background characteristics as the outcome variable. Overall, there are no significant changes in group composition: the p-value for the test of joint significance of all the characteristics is 0.861. When testing each characteristic separately, a couple demographic differences emerge (e.g., an increase in mother's birth-giving age); however, most variables remain balanced over the entire period. Importantly, there are no significant changes in characteristics that are strongly related to student achievement, such as parental income and education.

Other changes that coincide with the admission reform

A final crucial point is that there should be no other changes coinciding with the admission reform that would make the treatment and control group perform differently in school. To my knowledge, there are no other reforms occurring this year that would have an impact on student outcomes. However, around the time of the reform, there was a marked increase in school choice, with a growing number of compulsory school students opting out of the public school closest to their home (see the top panel of Figure A.1).

I perform several robustness checks to verify that changes in student sorting and school choice at the compulsory level do not affect my findings. Table A.3 reports the results. First, I show that the main estimates are not particularly sensitive to the inclusion of controls for teacher characteristics and peer characteristics (rows 4 and 5) or even time trends in these variables (rows 6 and 7). Next, in light of previous research indicating that independent schools

[20]High school admission centers report that it is not unusual for grades to increase significantly between the fall semester and the spring semester.

may be more likely to inflate grades, I show that my main results hold when restricting the sample to students who graduate from public schools (row 8).

When it comes to changes in school choice at the high school level, it may seem concerning that enrollment in independent schools starts to diverge in the treated and control regions around the time of the admission reform (see the bottom panel of Figure A.1). However, the gradual expansion of independent schools in the years prior to the reform can, in a sense, be considered a phase-in of the treatment that I study, given that independent schools have always used grades-based admissions. Likewise, the expansion of independent schools after the reform can also be considered part of the treatment: with increased choice, earning higher grades would now provide access to an even wider range of schools. Nevertheless, even though it is an imperfect control variable,[21] I show that including time trends in the share of high school students enrolled in independent schools at the time of application does not eliminate the treatment effect, though its magnitude drops to 7.6% of a standard deviation (row 9).

Another potentially concerning change during the study period was the adoption of a new grading scale (see section 2.1). The change in grading scale preceded the admission reform by two years, and all students were subject to the same change in rules. Figure A.2 suggests that the spread of the grade distribution was affected similarly in the treated and control regions. While these patterns are reassuring, the change of grading scale could still create differential trends in GPA if the initial shape of the GPA distribution differed in the treated and control regions. In particular, the anticipation effect that we observe in 1999 could be a delayed reaction to the new grading scale if there were more high-achieving students in the treated region and teachers were hesitant to set extreme grades in the first year after the adoption of the new grading standards.

To investigate this concern, I start by examining the share of observations in each decile of the national GPA distribution prior to the introduction of the new grading scale. Figure A.3 reveals that there are only minor differences, though there is slightly more density in the top two deciles of the distribution in the treated region. In order to rule out that these differences in density affect my results, I perform two robustness checks. First, I weigh observations in the treated region by the relative densities in each decile of the national distribution such that individuals in deciles that are overrepresented in the treated region get less weight in the regression and vice versa. Second, for each graduation cohort, I predict students' grades based on their background

[21] The increase in independent school enrollment may itself be an effect of the admission reform.

characteristics.[22] Then, I control for trends in predicted grades in the main difference-in-differences model. The results of these two checks are shown in rows 11 and 12 of Table A.3 respectively. Reassuringly, there is little change in the point estimates.

Lastly, a minor concern is that with the adoption of the new grading scale, schools stopped tracking students into different types of math and English classes. Thus, grades in math and English are not completely comparable over my entire study period. In row 13 of the table, I therefore show that my main estimates are unchanged if I exclude these subjects from the computation of the grade point average.

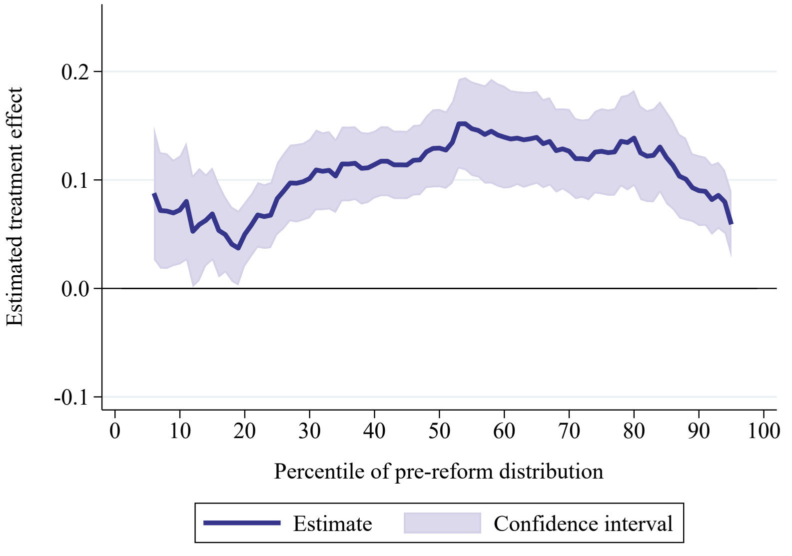
5.2 Distributional effects of the reform

Thus far, my analysis has focused on the average effect of the admission reform. However, the size of the effect could differ across the grade distribution. For example, high-achieving students may have weak responses to the reform if their current level of effort puts them close to the maximum GPA or safely above the predicted admission cutoff for their desired school and program. Similarly, low-achieving students may have weak responses given their low likelihood of being admitted to prestigious schools even if they exert a lot of effort to improve their grades. By contrast, average-achieving students are on the margin of admission to many popular programs, and with a reasonable amount of effort, they can raise their GPA above the expected admission cutoffs. Moreover, after the end of proximity-based admissions, average-achievers who live near prestigious schools face the greatest risk of being crowded out by high-achievers who live far away; under the new rules, they must earn higher grades to stand a chance of admission. Thus, it is reasonable to expect stronger responses in the middle of the distribution.

To investigate this possibility, I study the distributional effects of the reform using the unconditional quantile regression method that I described in section 3. The solid blue line in Figure 4 plots the estimates of the unconditional quantile treatment effects at each percentile of the pre-reform treatment distribution, and the shaded region depicts the 95% confidence interval for the point estimates. I also report the results at several key points of the distribution in Table A.4. Notably, there are sizable and significant effects across the entire grade distribution, indicating that students of all abilities improved their performance as a result of the reform. However, the effect size is not uniform. The quantile treatment effects roughly follow an inverted U-shaped pattern,

[22] I use the national grade distribution, but exclude the treated region (Stockholm municipality) when predicting grades.

Figure 4. Distributional effects of the reform.



Notes: The solid blue line traces out the estimated quantile treatment effect from equation (3) at each percentile of the pre-reform treatment distribution. The shaded region shows the 95% confidence interval for the estimates.

with smaller effects in the bottom and top of the distribution. The peak occurs just above the median, where the estimated treatment effect is about 15% of a standard deviation.[23]

The pattern of non-linear effects is consistent with a story in which student responses drive the increases in grades. The effects are strongest in the middle of the distribution, where the majority of cutoffs to academic programs are and students thus have the most to gain from increasing their grades (see Figure A.6). However, this pattern could also be observed if grade inflation drives the results. For example, parents of students who are on the margin of admission to a prestigious school may be more likely to pressure teachers into raising their children's grades. Thus, I spend the next two sections trying to disentangle the grade inflation channel from changes in student effort.

[23] A potential concern is the extent to which the change of grading scale in 1998 affected the shape of the GPA distribution in the two groups. In Figure A.4, I show that the results are largely similar if I restrict the sample to cohorts 1998 and later, all of whom graduated with the same grading scale. An exception is the very bottom of the distribution, where the estimates are quite large in magnitude but also very noisy. This is likely due to the fact that there is very little density in this portion of the grade distribution, and the estimates are highly sensitive to the estimates of the kernel density function (see Figure A.5).

6. Direct checks of the grade inflation channel

A crucial question is whether behavioral responses from students or teachers drive the effects that I observe. In either case, the fact that grades increased due to the reform has meaningful consequences at the individual level. However, from a policy perspective, the desire is to increase effort and knowledge, not to alter grading practices. If the reform increased grade inflation and widened the differences in grading standards between regions, this could have negative welfare effects (Nordin et al., 2019) and ultimately even reduce student effort and learning (Betts and Grogger, 2003).

In this section, I use data on standardized exams to investigate whether grade inflation drives my results. As discussed earlier, there are no external evaluators on these exams; rather, teachers grade their own students according to national grading criteria. Two features of the grading criteria are ideal for assessing the extent to which teachers inflate grades. First, the standardized exams in math are graded on a numerical scale with pre-determined cutoffs for each letter grade. Second, teachers in math, English, and Swedish are supposed to take exam grades into account when setting final course grades. In section 6.1, I exploit the first feature to gauge the degree of grade inflation in treatment versus control regions after the reform. In section 6.2, I exploit the second feature to assess the extent to which grading standards became more lenient due to the reform.

6.1 Test score manipulation on maths exams

Since the end of my study period, all ninth grade students are required to take standardized tests in math. These exams are graded on a numeric scale from 0-70 with specific cutoffs corresponding to a grade of Pass (*G*) and Pass with Distinction (*VG*).[24] The grade cutoffs vary from year to year, and there is some teacher discretion in how the grading criteria are applied. Thus, it is unlikely that students can precisely manipulate their score and strategically sort just above the cutoff required for a certain grade. This implies that the test score distribution should be continuous around the cutoffs for the grades of Pass and Pass with Distinction. If there is a significant bunching of scores just above these cutoffs, it must be because teachers manipulate (or "inflate") scores to push students from a lower to a higher grade.

As a first check on the grade inflation channel, I measure the amount of bunching above each cutoff in the test score distribution separately by treated

[24] There is no specific cutoff for a grade of Pass with Highest Distinction (*MVG*), but to obtain MVG, the student must pass the VG threshold.

and control regions. While the extent to which teachers manipulate scores on the maths exam does not perfectly correspond to how much they inflate final course grades, it is nevertheless informative about how lenient teachers are in their grading practices. If grade inflation drives my result, it is likely that teachers in the treated region manipulate test scores to a larger extent than teachers in the control region; in that case, we should see a greater degree of bunching above the thresholds.

To test this hypothesis, I implement a bunching estimator similar in spirit to the one in Dee et al. (2019) and inspired by the methodological approach in Chetty et al. (2011). I start by plotting the observed distribution of test scores. By visual inspection, I determine the range of scores above each cutoff where there appears to be excess mass, as well as the range of scores below each cutoff where there appears to be missing mass (i.e., the region where the excess mass would have been if teachers did not manipulate scores). Then, I estimate the counterfactual distribution that would have occurred in a world without test score manipulation by fitting a flexible polynomial to the observed distribution,[25] excluding points in the region of excess mass. I use an iterative procedure to shift the counterfactual distribution in the region of missing mass upwards until the estimated missing mass equals the estimated excess mass. To obtain the standard errors, I use the bootstrap procedure described in Chetty et al. (2011).

Figure 5 shows the results of the bunching analysis for the 2004 graduation cohort. While test score manipulation is evident in both regions—at least around the threshold to pass—the difference in the degree of manipulation between the treated and control regions is quite negligible. The total amount of manipulation is only marginally higher in the treated region than the control region (1.587% versus 1.478%). Assuming that grading practices on the maths exam are reflective of grading practices in general, these results provide little to no evidence that grade inflation is stronger in the treated region than the control region after the admission reform.

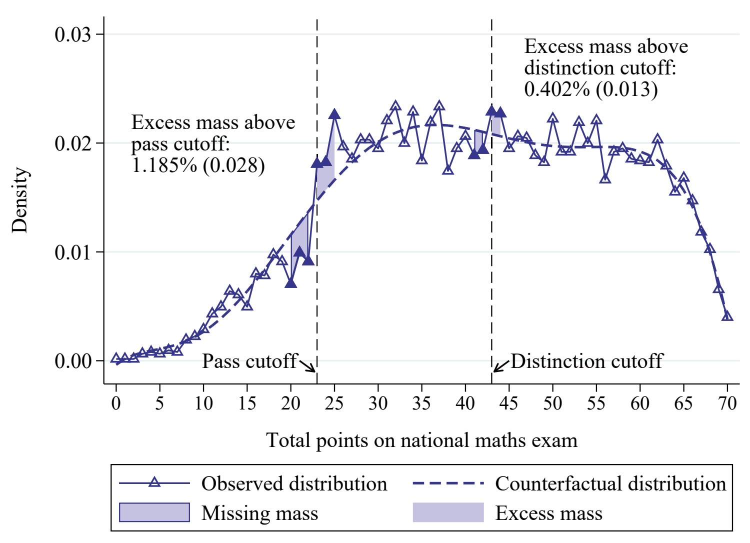
6.2 Correspondence between exam grades and course grades

According to national guidelines, teachers in math, English, and Swedish are supposed to take students' standardized exam grades into account when setting final course grades. However, they can exercise discretion and assign either a higher or a lower mark. I use this fact to measure how lenient teachers are

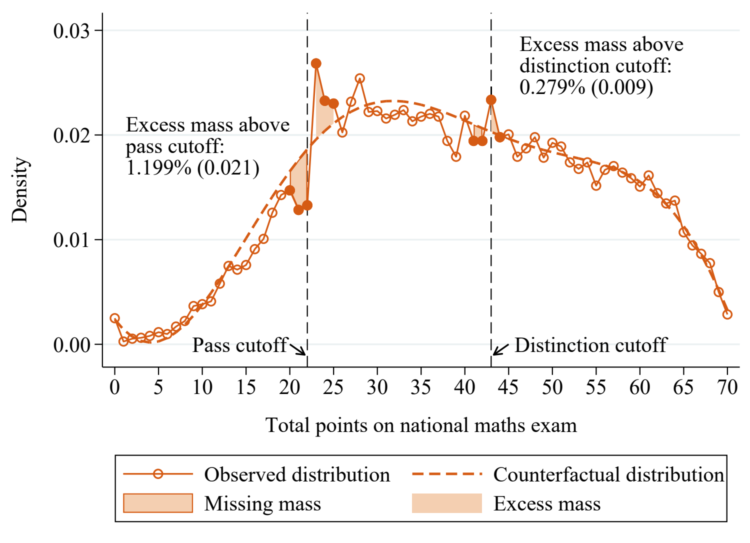
[25] I allow the order of the polynomial to vary by region and year. I use the Akaike information criterion (AIC) to guide my choice. The results are not sensitive to changing the order of polynomial.

Figure 5. Bunching in the national maths test distribution (2004 cohort).

(a) Treated region



(b) Control region



Notes: In this figure, I plot the distribution of test scores on the ninth grade standardized mathematics exam from spring 2004 separately by treated and control region. Each triangle and circle corresponds to the fraction of test takers who earned the given score on the horizontal axis. The vertical dashed lines show the cutoffs for grades of Pass and Pass with Distinction. The curved dashed line depicts the counterfactual distribution that would have occurred if teachers did not manipulate scores around these cutoffs. The degree of manipulation at each cutoff is given by the estimated excess mass, with standard errors from a parametric bootstrap procedure shown in parentheses.

when setting grades. For each student i, I construct an “inflation index” that measures the extent to which their course grade in subject c is inflated over their exam grade in that subject. I first convert students’ grades to a numeric scale,[26] and then take the difference between their course grade and exam grade:

$$InflationIndex_{i,c} = CourseGrade_{i,c} - ExamGrade_{i,c}$$

Next, using data from the three municipalities in Stockholm county for which I have both pre- and post-reform data on grades and standardized tests, I estimate my main difference-in-differences model using the inflation index as the dependent variable.[27] Table 3 shows the results. In column 1, I begin by verifying that the change in GPA for the subsample is in line with my main results for all of Stockholm county. The estimated effect on standardized GPA is 0.14, which is notably larger than the main point estimate of 0.10 from Table 2. However, due to data limitations, I only include cohorts 1998 and 2004 in the current estimation. According to the event study plot in Figure 3, the annual estimate relative to 1998 is 0.137 for the 2004 cohort, which is in line with the result for the current subsample.

Reassured that the results for the subsample are similar to the results for the main sample, I proceed with an analysis of the inflation index. The reported coefficient on $Treat \times Post$ should be positive and significantly different from zero if the admission reform increased teachers’ grading leniency. Column 2 reports the results for math and indicates that, if anything, grading leniency actually decreased in the treated region compared to the control region. Meanwhile, the point estimates for English in column 3 and Swedish in column 4 are small in magnitude and statistically indistinguishable from zero, indicating that there was no change in grading leniency in these subjects due to the reform. Altogether, these results suggest that the grade inflation channel is not driving my main results.

[26] I assign each grade the number of points that it is worth in the official calculation of students’ GPA, i.e., 0 points if the student does not pass; 10 points for pass; 15 points for pass with distinction; and 20 points for pass with highest distinction. These numbers are set by national regulations and cannot vary across regions.

[27] In this case, I assign students to treated and control regions based on where they graduate from compulsory school. I am required to change my method of treatment assignment due to the fact that I have different personal identifiers for the standardized test data from 1998 and my main analysis data from Statistics Sweden, and I do not observe municipality of residence or pre-reform municipality of residence in the standardized test data from 1998.

Table 3. Effect of the admission reform on the grade inflation index.

		Grade inflation index in:		
	GPA	Math	English	Swedish
	(1)	(2)	(3)	(4)
Treat × Post	0.140	-0.868	0.026	-0.099
	(0.054)***	(0.481)*	(0.174)	(0.362)
Pre-reform mean	0.256	1.420	-0.211	0.866
Observations	13,396	11,502	13,396	12,041

Notes: Each entry reports the point estimate of the coefficient on *Treat* × *Post* from a different difference-in-differences model. In column 1, the dependent variable is standardized GPA. In columns 2 to 4, the dependent variable is the inflation index in either math, English, or Swedish. All models are estimated using the three municipalities for which I have pre-reform and post-reform standardized test data. The pre-reform mean refers to the average GPA in Stockholm municipality in 1998; it is notably higher than the average GPA for all of Stockholm municipality that year due to selection of students for the pilot exam. I report robust standard errors in parentheses and denote the significance level of the estimate with stars: *** for p-value < 0.01; ** for $p < 0.05$; and * for $p < 0.10$.

7. Corroborative evidence of the mechanisms

To corroborate the evidence from the previous section, I perform two heterogeneity analyses. First, I test whether the admission reform had a larger effect on students who graduated from schools that arguably had a greater incentive to inflate grades. Second, I test whether the reform had a stronger effect on students who arguably had a greater incentive to boost their grades.

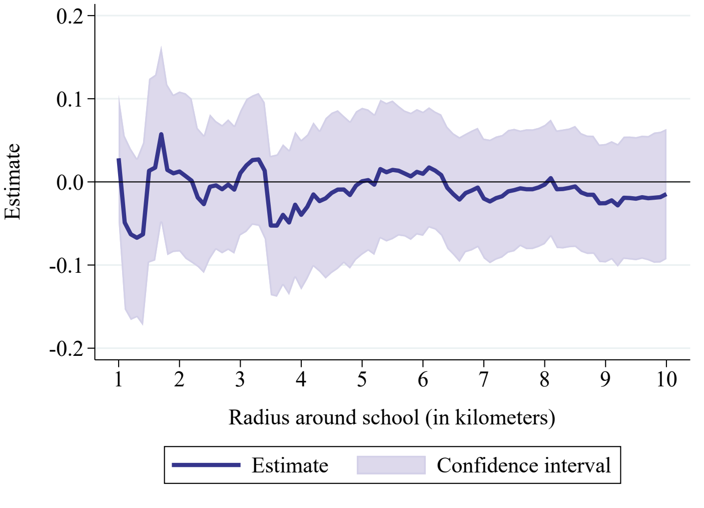
7.1 Effects by degree of compulsory school competition

In Sweden, there is a high degree of school choice. Schools that face more competition are at a higher risk of losing students and the per-pupil voucher that the government provides to whatever school they attend. Thus, they have a stronger incentive to boost grades in an attempt to market themselves as a high-performing school. While these schools may have always used grades as a way to attract students, the reform increased the marketing value of high grades. Thus, if grade inflation is driving the results, it would be reasonable to expect stronger increases in the grades of students attending schools in higher-competition areas.

In order to check this, I perform the main difference-in-differences analysis by degree of compulsory school competition. I create a variable, $Competition_i$, that measures the amount of competition faced by the compulsory school that student i graduates from. For each school, I measure competition as the fraction of grade nine enrollment covered by other schools within an x kilometer radius; thus, the measure equals zero if there are no other schools within an x kilometer radius and grows closer to one the smaller a school's enrollment share is.[28]

To measure heterogeneous effects by competition, I fully interact the *Competition* variable with the right-hand side of the main difference-in-differences model in equation (1). The coefficient of interest is the estimate on triple interaction term, $Competition \times Treat \times Post$, and it would be positive and significant if the reform had larger effects on schools facing higher competition. It is unclear exactly what radius is relevant for measuring school competition, so I try out different radii from one to 10 kilometers in 0.1 km increments. I plot the estimates for different radii in Figure 6. The majority of the estimates are negative in sign and statistically indistinguishable from zero, indicating that the reform did not have differential effects by level of compulsory school competition. Because the incentives to inflate grades should vary with the degree of competition, this corroborates the evidence from section 6, which suggested that grade inflation is not driving the results.

Figure 6. Effects by degree of compulsory school competition within x kilometers.



Notes: The solid blue line plots the estimate of the triple interaction term $Competition \times Treat \times Post$ when defining school competition within different distances from the school. The shaded region shows the 95% confidence interval for the estimates.

[28]The competition measure is calculated as the average level of competition during the pre-reform period, but the results are almost identical if I allow the measure to vary year by year.

The finding that the effect of the admission reform does not vary by degree of competition also speaks against another possible behavioral response on the school side: i.e., that schools might improve their quality in response to the reform. Just as the pressure to inflate grades is stronger for schools subject to more competition, so too is the pressure to raise their quality. Thus, it would be reasonable to expect larger effects in schools subject to more competition if schools reacted to the reform by improving their effectiveness. As discussed above, there is no evidence that this is the case.

7.2 Effects by likelihood to enroll in an academic track

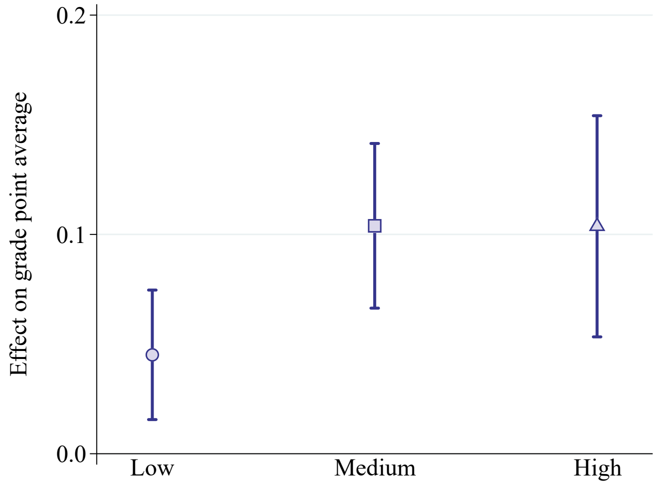
During my period of study, the most popular programs in Stockholm were the social sciences and natural sciences program.[29] These programs were also offered at the most prestigious schools. Thus, the reform created the greatest shift in incentives for students who wanted to attend these tracks. If changes in student behavior drive the results, we might expect to see stronger results for students who want to attend these programs.

As a check on the student effort channel, I investigate whether the reform had heterogeneous effects by students' predicted probability of applying to the social sciences or natural sciences track in the top choice on their high school application. I use their predicted probability of applying to these tracks rather than their observed preferences, because it is likely that the program they apply to is endogenous to the reform. With pre-reform data from years 1994 to 1998, I obtain the predicted probability for each student via logistic regression. The vector of covariates includes gender, age, foreign background, parents' income, parents' years of schooling, and parents' field of education.

Figure A.7 in the appendix shows the histogram of predicted probabilities by treated and control region. I divide students into three groups: low likelihood (bottom third of the predicted probability distribution), medium likelihood (middle third of the distribution), and high likelihood (top third of the distribution) of applying to social or natural sciences as a top choice. Then, I estimate the main differences-in-differences model separately for each of the three groups. The results are shown in Figure 7. Although the effects are a bit imprecise, the pattern follows what we would expect if student responses drive the results: the effects are smaller for the low likelihood group relative to the medium and high likelihood group.

[29] I consider the technology program a natural sciences program: it was officially classified as one in the pre-reform period, but in the post-reform period, it was a standalone program.

Figure 7. Effects by likelihood of preferring an academic track.



Notes: This figure shows the difference-in-differences estimate, i.e., λ from equation (1), when running the analysis separately for students with a low, medium and high likelihood of listing an academic track (excluding the arts program) in the top choice on their high school application. The three groups are determined using the first, second, and third quantile of the distribution of predicted probabilities from a logistic regression (see Figure A.7 in the appendix).

8. Longer-run effects on student achievement

The evidence indicates that the admission reform raised compulsory school grades due to an increase student effort rather than an increase in teachers' grading leniency. While this suggests that performance incentives could be an effective policy lever for raising student achievement, it is relevant to know whether the effects persist over time. Thus, I repeat the main difference-in-differences analysis using various high school outcomes as the dependent variable. In particular, I evaluate the effect of the reform on the probability of earning a high school diploma and on grades in core high school courses that all students must complete regardless of the program that they enroll in.

Table 4 presents the results. Before estimating the effects on high school performance, I show that the admission reform did not have an impact on selection into high school. Column 1 reports the estimated effect on the probability of enrollment. The point estimate is equal to zero and statistically insignificant, indicating that the reform did not affect attendance patterns. However, the reform did have a positive effect on the probability of graduation. Turning to column 2, the difference-in-differences estimate indicates around a 2.4 percentage point increase in the likelihood of earning a high school diploma, which is a 3.4% increase relative to the baseline probability. Panel (a) of Fig-

Table 4. Effect of the admission reform on high school outcomes.

	Attend (1)	Graduate (2)	Core GPA (3)
Treat × Post	0.000	0.024	-0.001
	(0.001)	(0.007)***	(0.012)
Wild bootstrap p-value	0.845	0.141	0.941
Pre-reform mean in the treated region	0.974	0.712	0.234
Observations	178,116	178,116	129,845

Notes: Each column reports the difference-in-differences estimate, i.e., λ from equation (1), using various high school outcomes as the dependent variable. In column 1, the dependent variable is a binary indicator equal to one if the individual enrolls in high school the year they exit grade nine. In column 2, it is a binary indicator equal to one if the individual graduates from high school within five years of exiting grade nine. In column 3, it is an individual's average GPA in core high school courses (i.e., courses taken by all individuals), standardized with respect to the national distribution for their graduation cohort. High school grades are only available for graduates.

ure A.8 traces out the dynamics of this effect over time. Reassuringly, the pattern is relatively similar to the pattern observed for compulsory school grades, with an immediate effect on the cohort who exited ninth grade during the anticipation period and a slightly larger effect on the cohort who exited ninth grade immediately after the reform was implemented. The fact that the likelihood of graduation increased due to the reform suggests that the achievement gains may persist in the medium-run. On the other hand, column 3 shows that conditional on graduation, the reform did not have an impact on high school grades, which is consistent with the finding of an earlier study by Söderström (2006) that estimated the effects using only one year of post-reform data. However, this last result should be interpreted with caution: the fact that grades are unobserved for all students who fail to earn a high school diploma is concerning given that the reform affected the probability of completion.

9. Concluding remarks

The results of this paper suggest that strengthening the performance incentives implicit in the design of the education system can have a positive effect on student achievement. I provide evidence from a high school admission reform that replaced proximity-based school assignment with strictly grades-based admissions, thereby strengthening the incentives for compulsory students to earn high grades. My difference-in-differences analysis shows that compulsory school grades increased due to the reform, with positive effects along the entire grade distribution and the largest shifts around the median.

While this is not the first study to find that grades change in response to grades-based incentives, I contribute to the literature with a rigorous investigation of the mechanisms driving the findings. The checks that I perform all render support for the hypothesis that the grade increases are driven by student responses, rather than more lenient grading practices. This finding is important from a policy perspective, because the aim of performance incentives is to boost student effort and enhance learning. If the effect had instead been the result of grade inflation, this could have undesirable welfare effects (Nordin et al., 2019) and even lower student motivation and effort (Betts and Grogger, 2003). However, given that I find no evidence that grade inflation drives my effect, my results suggest that altering the performance incentives in the education system may be an effective policy tool to raise student achievement. An open question is how long these effects persist and whether they translate into longer-term achievement gains.

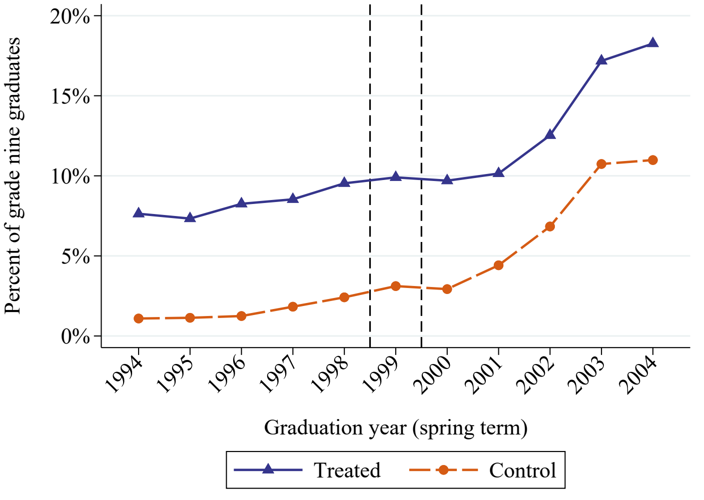
Another important conclusion from my analysis is that a student's achievement in a given grade level is sometimes endogenous to features of the educational system that the student will experience at later grade levels. From an econometric perspective, this implies that we should be careful when including prior student achievement in models evaluating the effect of policies that change the education system at later grades. Moreover, when evaluating educational interventions, policymakers should consider its potential effects on students of all ages, not just students in the targeted grade levels.

Appendix

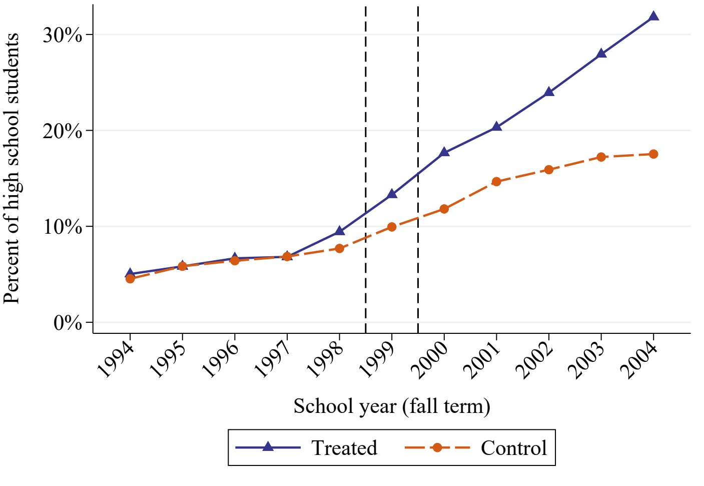
Figures

Figure A.1. Share of students in independent schools.

(a) Grade nine graduates

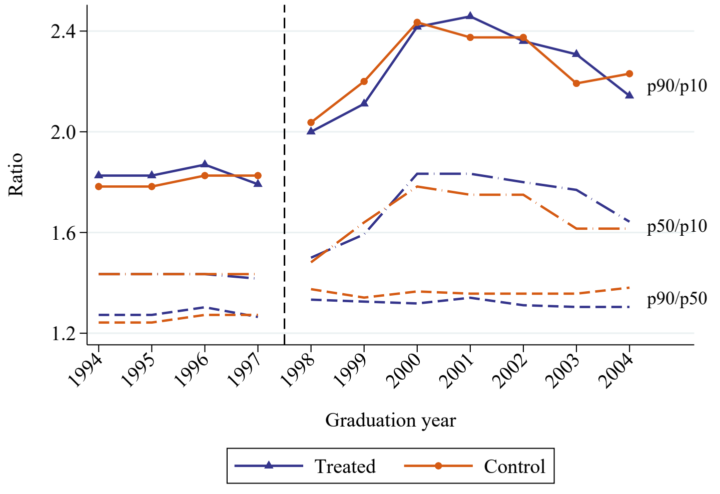


(b) High school students



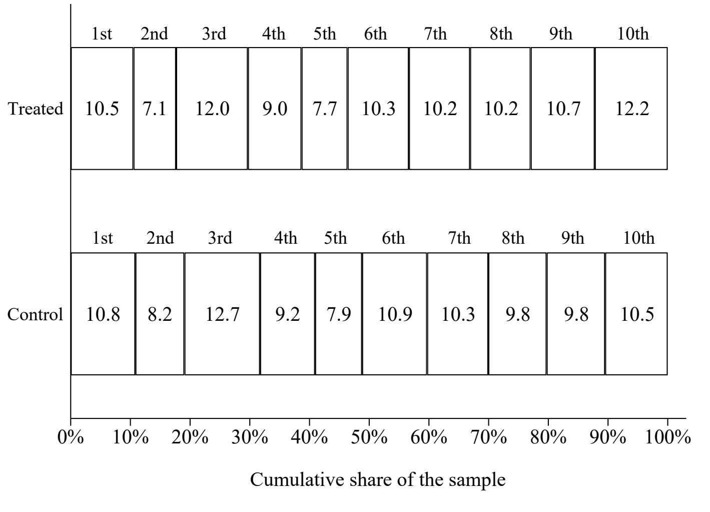
Notes: Panel (a) plots the percent of students who graduate from grade nine at an independent school. Panel (b) plots the percent of students who attend an independent high school, excluding a small number of schools without full high school programs (e.g., art schools with one- or two-year programs in painting). The treated group includes students enrolled in Stockholm municipality, and the control group includes students enrolled in the rest of Stockholm county.

Figure A.2. Spread in the grade distribution over time.



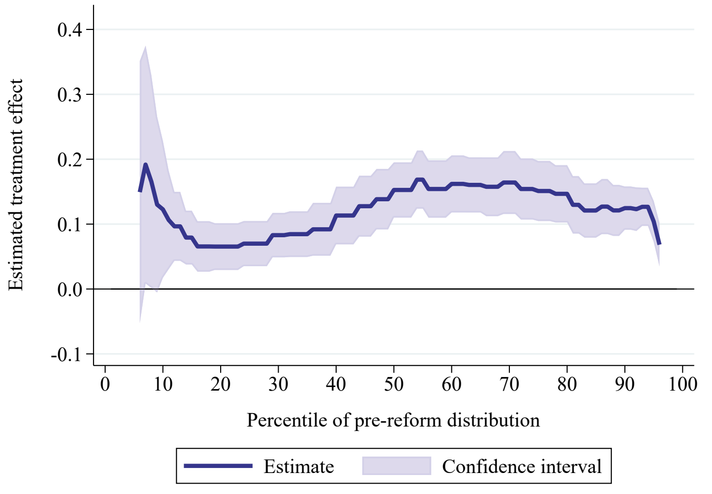
Notes: The blue lines plot ratios for the treated region, and the orange lines plot ratios for the control region. The dashed vertical line indicates the switch from a relative grading scale to criterion-based grading.

Figure A.3. Densities in different parts of the GPA distribution prior to the introduction of criterion-based grading.



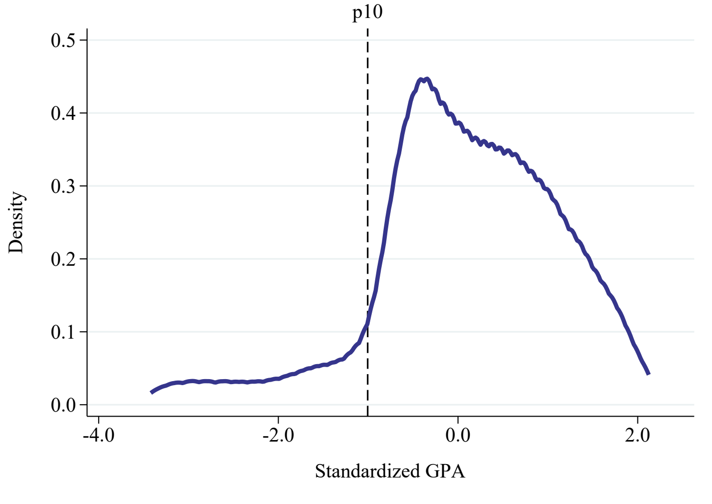
Notes: For both the treated and control group, this figure reports the share of observations in each decile of the national GPA distribution prior to the introduction of criterion-based grading in 1998. For example, 7.7% of the observations in the treated group fall in the 5th decile of the national GPA distribution, while the corresponding number for the control group is 7.9%.

Figure A.4. Distributional effects on standardized GPA (restricted sample).



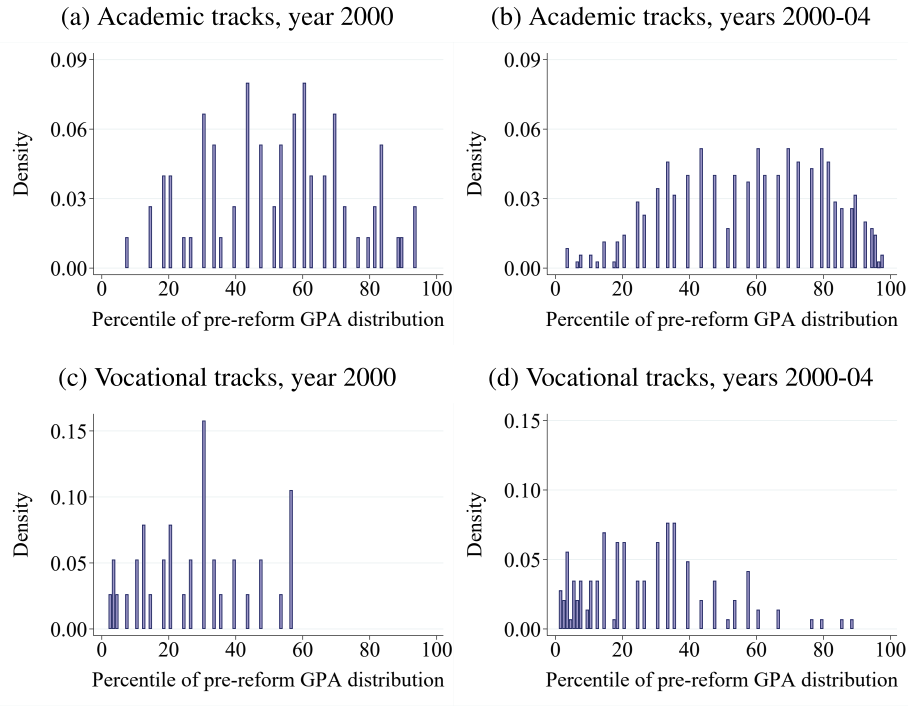
Notes: Similar to Figure 4 in the main paper, this figure plots the unconditional quantile treatment effects at different percentiles of the pre-reform GPA distribution. However, the sample is restricted to individuals who graduated from ninth grade between 1998 and 2004, i.e., after the switch from relative to criterion-referenced grading.

Figure A.5. Kernel density plot of the pre-reform grade distribution.



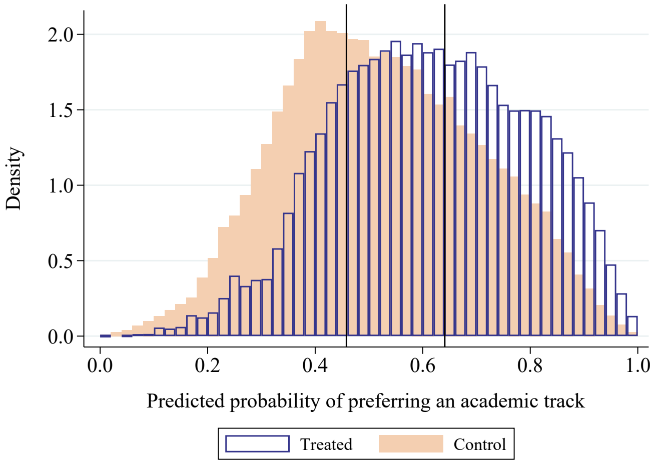
Notes: The dashed vertical line marks the 10^{th} percentile of the pre-reform grade distribution in the treated region.

Figure A.6. Post-reform distribution of admission cutoffs for programs offered at public schools in the treated region.



Notes: These figures show the distribution of admission cutoffs for academic tracks (excluding the Arts program, which has special admission criteria) and vocational tracks at public schools in the treated region. The left-hand panels show the cutoffs in the first post-reform year and the right-hand panels show the cutoffs for all post-reform years in the study period. Programs without binding admission cutoffs are excluded from the histograms. All cutoffs are converted to their percentile of the pre-reform treatment distribution.

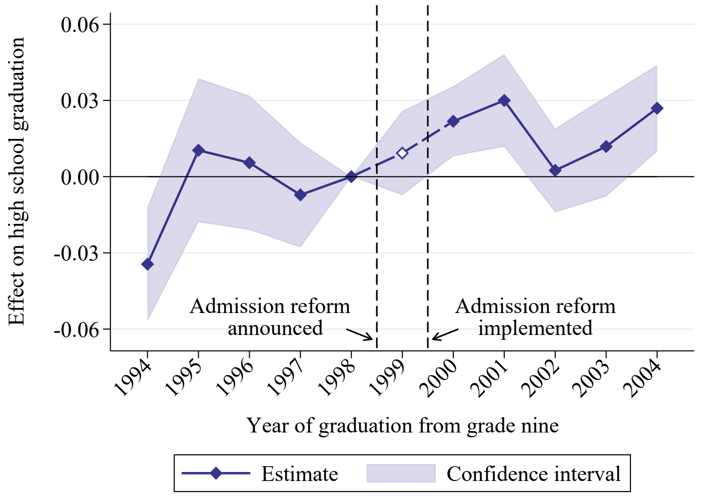
Figure A.7. Histogram of predicted probability of preferring an academic track.



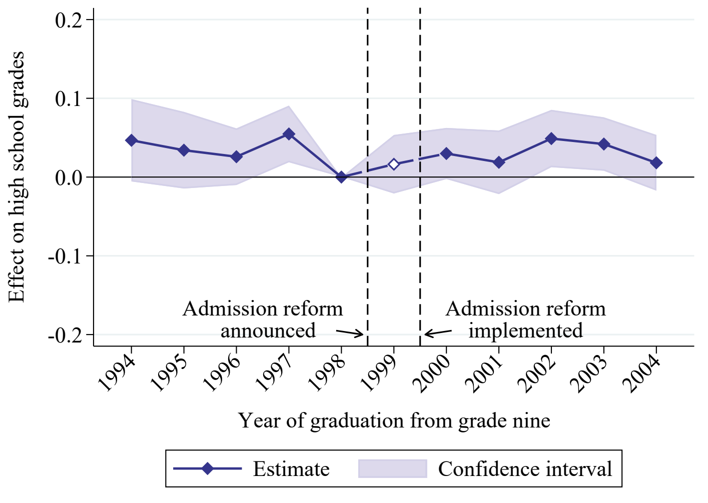
Notes: This figure reports the distribution of grade nine graduates' predicted likelihood to list an academic track (other than the arts track) in the top choice on their high school application. The white bars with blue outlines correspond to the treated region, and the orange bars correspond to the control region.

Figure A.8. Event study plots for high school outcomes.

(a) Probability of graduating from high school



(b) Standardized GPA in core high school courses



Notes: Each panel plots the estimates from an event study specification of equation (1), in which I replace the post-reform indicator with a set of time dummies. I omit the time dummy for 1998, such that all annual coefficients are measured relative to the year prior to the announcement of the reform. In panel (a), the dependent variable equals one if an individual graduates within five years of exiting grade nine. In panel (b), the dependent variable is standardized GPA in core high school courses (i.e., courses taken by all students regardless of the high school track they enroll in).

Tables

Table A.1. Descriptive statistics.

	All grade 9 graduates	Treated group	Control group
Girl	0.489	0.482	0.485
Age	15.020	15.039	15.037
	(0.243)	(0.198)	(0.191)
Foreign background	0.138	0.190	0.171
Mother's birth-giving age	28.012	29.147	28.483
	(5.119)	(5.232)	(5.217)
Number of siblings	1.814	1.557	1.734
	(1.201)	(1.118)	(1.126)
Parents are married	0.630	0.522	0.597
Parents' income	374.718	449.324	443.738
	(239.476)	(350.691)	(318.589)
Mother's employment status	0.814	0.833	0.846
Father's employment status	0.851	0.823	0.859
Mother's years of schooling	11.630	12.225	11.794
	(2.512)	(2.743)	(2.605)
Father's years of schooling	11.458	12.356	11.828
	(2.864)	(3.019)	(2.880)
Mother has an elite degree	0.023	0.052	0.036
Father has an elite degree	0.072	0.140	0.112
Missing father's information	0.032	0.024	0.016
Attended public school	0.976	0.942	0.955
Grade 9 GPA (standardized)	0.002	0.178	0.051
Observations	1,103,387	58,541	119,575

Notes: The first column reports the average characteristics for all grade nine graduates in Sweden between 1994 and 2004. The treated group refers to graduates whose mother lived in Stockholm municipality at the end of 1993, and the control group refers to graduates whose mother lived in another municipality in Stockholm county at the end of 1993.

Table A.2. Balance of covariates.

	Estimate of Treat × Post	Clustered Std. Err.	Bootstrap P-value
Girl	-0.002	0.003	0.537
Age	0.003	0.002	0.300
Foreign background	0.001	0.003	0.626
Mother's birth-giving age	0.556	0.051	0.000
Number of siblings	-0.020	0.011	0.285
Parents are married	0.036	0.004	0.011
Parents' income	6.026	6.327	0.468
Mother's employment status	0.000	0.004	0.931
Father's employment status	0.003	0.003	0.441
Mother's years of schooling	-0.004	0.027	0.899
Father's years of schooling	0.060	0.036	0.307
Mother has an elite degree	0.004	0.003	0.360
Father has an elite degree	0.006	0.003	0.309
Missing father's information	-0.002	0.001	0.284
Attended public school	0.006	0.008	0.506
P-value for test of joint significance: 0.861			

Notes: Each entry in the second column reports the estimate of the interaction term in equation (1) from separate regressions in which I use the characteristic in the first column as the dependent variable. The third column reports standard errors robust to clustering at the municipal level. The fourth column reports the p-value when I perform inference using the wild-cluster bootstrap procedure suggested by Cameron et al. (2008).

Table A.3. Robustness checks for main difference-in-differences analysis.

	Estimate of Treat × Post	Clustered standard error	Wild bootstrap p-value
1. Baseline estimation, column 2 from Table 2	0.100	0.015	0.034
2. Assigning treatment by graduation municipality	0.097	0.018	0.056
3. Assigning treatment by current residence	0.107	0.015	0.042
4. Including controls for teacher characteristics	0.097	0.016	0.029
5. Including controls for peer characteristics	0.103	0.015	0.037
6. Controlling for trends in teacher characteristics	0.095	0.017	0.072
7. Controlling for trends in peer characteristics	0.087	0.018	0.101
8. Sample restricted to public schools	0.109	0.016	0.035
9. Controlling for trends in share of students who attend independent high schools	0.076	0.223	0.126
10. Sample restricted to cohorts with criterion-based grading	0.115	0.020	0.049
11. Weighing observations to adjust for differences in pre-reform densities	0.101	0.015	0.033
12. Controlling for trends in predicted grades	0.091	0.014	0.040
13. Excluding Math and English from grade point average	0.100	0.017	0.047
14. Including school fixed effects	0.101	0.019	0.122

Notes: Each row corresponds to a different robustness check, with the exception of the first row, which repeats the main estimate for comparison. The second column reports the estimate of the interaction term in equation (1) for each robustness check. The third column reports standard errors clustered at the municipal level. The fourth column reports the p-value when I perform inference using the wild-cluster bootstrap procedure suggested by Cameron et al. (2008). In row 2, teacher characteristics include school-level averages of teachers' certification status, experience, and contract type (part-time or permanent). In row 3, peer characteristics include school-level averages of parents' income and years of education, as well as the share of girls and students with a foreign background.

Table A.4. Distributional effects of the admission reform.

	10th perc. (1)	25th perc. (2)	50th perc. (3)	75th perc. (4)	90th perc. (5)
Treat × Post	0.072	0.083	0.129	0.126	0.090
	(0.024)***	(0.016)***	(0.017)***	(0.019)***	(0.016)***
Wild bootstrap p-value	0.181	0.057	0.028	0.047	0.056
Pre-reform level	-1.176	-0.469	0.136	0.852	1.426
Observations	178,116	178,116	178,116	178,116	178,116

Notes: This table reports the unconditional quantile treatment effects at different percentiles of the pre-reform grade distribution in the treated region. Standard errors are clustered at the municipal level and reported in parentheses. Stars denote significance levels: *** for p-value < 0.01; ** for $p < 0.05$; and * for $p < 0.10$. Because there are only 25 clusters, I also report the p-value when I perform inference using the wild-cluster bootstrap procedure suggested by Cameron et al. (2008).

References

Angrist, J., Lang, D., and Oreopoulos, P. (2009). Incentives and Services for College Achievement: Evidence from a Randomized Trial. *American Economic Journal: Applied Economics*, 1(1):136–163.

Angrist, J. D., Lavy, V., Leder-Luis, J., and Shany, A. (2019). Maimonides' Rule Redux. *American Economic Review: Insights*, 1(3):309–324.

Baumert, J. and Demmrich, A. (2001). Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education*, 16(3):441–462.

Bettinger, E. P. (2011). Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores. *The Review of Economics and Statistics*, 94(3):686–698.

Betts, J. R. and Grogger, J. (2003). The impact of grading standards on student achievement, educational attainment, and entry-level earnings. *Economics of Education Review*, 22(4):343–352.

Black, S. (1999). Do Better Schools Matter? Parental Valuation of Elementary Education. *The Quarterly Journal of Economics*, 114(2):577–599.

Burgess, S., Greaves, E., and Vignoles, A. (2019). School choice in England: Evidence from national administrative data. *Oxford Review of Education*, 45(5):690–710.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3):414–427.

Chetty, R., Friedman, J. N., Olsen, T., and Pistaferri, L. (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *The Quarterly Journal of Economics*, 126(2):749–804.

Dee, T. S., Dobbie, W., Jacob, B. A., and Rockoff, J. (2019). The Causes and Consequences of Test Score Manipulation: Evidence from the New York Regents Examinations. *American Economic Journal: Applied Economics*, 11(3):382–423.

Diamond, R. and Persson, P. (2016). The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests. Technical Report w22207, National Bureau of Economic Research, Cambridge, MA.

Fajnzylber, E., Lara, B., and León, T. (2019). Increased learning or GPA inflation? Evidence from GPA-based university admission in Chile. *Economics of Education Review*, 72:147–165.

Figlio, D. N. and Lucas, M. E. (2004). What's in a Grade? School Report Cards and the Housing Market. *American Economic Review*, 94(3):591–604.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, 77(3):953–973.

Fredriksson, P., Öckert, B., and Oosterbeek, H. (2013). Long-Term Effects of Class Size. *The Quarterly Journal of Economics*, 128(1):249–285.

Gneezy, U., Meier, S., and Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, 25(4):191–210.

Grove, W. A. and Wasserman, T. (2006). Incentives and Student Learning: A Natural Experiment with Economics Problem Sets. *The American Economic Review*, 96(2):447–452.

Haraldsvik, M. (2014). Does Performance-Based Admission Incentivize Students? Unpublished manuscript.

Havnes, T. and Mogstad, M. (2015). Is universal child care leveling the playing field? *Journal of Public Economics*, 127:100–114.

Hedges, L. V., Pigott, T. D., Polanin, J. R., Ryan, A. M., Tocci, C., and Williams, R. T. (2016). The Question of School Resources and Student Achievement: A History and Reconsideration. *Review of Research in Education*, 40:143–168.

Hvidman, U. and Sievertsen, H. H. (2019). High-Stakes Grades and Student Behavior. *Journal of Human Resources*, pages 0718–9620R2.

Jackson, C. K. (2020). Does School Spending Matter? The New Literature on an Old Question. In *Confronting Inequality: How Policies and Practices Shape Children's Opportunities*, pages 165–186.

Jackson, C. K., Johnson, R. C., and Persico, C. (2016). The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. *The Quarterly Journal of Economics*, 131(1):157–218.

Jalava, N., Joensen, J. S., and Pellas, E. (2015). Grades and rank: Impacts of non-financial incentives on test performance. *Journal of Economic Behavior & Organization*, 115:161–196.

Koerselman, K. (2013). Incentives from curriculum tracking. *Economics of Education Review*, 32:140–150.

Krueger, A. B. and Whitmore, D. M. (2001). The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR. *The Economic Journal*, 111(468):1–28.

Levitt, S. D., List, J. A., Neckermann, S., and Sadoff, S. (2016). The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance. *American Economic Journal: Economic Policy*, 8(4):183–219.

Metcalfe, R., Burgess, S., and Proud, S. (2019). Students' effort and educational achievement: Using the timing of the World Cup to vary the value of leisure. *Journal of Public Economics*, 172:111–126.

Molin, E. (2019). *Through the Ranks: Essays on Inequality, Status and School Choice*. PhD thesis, Stockholm School of Economics.

Nordin, M., Heckley, G., and Gerdtham, U. (2019). The impact of grade inflation on higher education enrolment and earnings. *Economics of Education Review*, 73:101936.

Orrenius, A. (1997). Budgetdebatten: Gymnasieantagning ska utredas [The Budget Debate - High school admissions to be investigated]. *Dagens Nyheter*. Published November 29, 1997.

Skolverket (2005). National assessment and grading in the Swedish school system. Technical report, Skolverket, Stockholm.

Skolverket (2007). Provbetyg-Slutbetyg-Likvärdig bedömning? En statistik analys av sambandet mellan nationella prov och slutbetyg i grundskolan, 1998-2006. [Exam scores - Final grades - Equivalent assessment? A statsitical analysis of the relationship between national exams and final grades in compulsory school, 1998-2006]. Technical report, Skolverket, Stockholm.

Stinebrickner, R. and Stinebrickner, T. R. (2008). The Causal Effect of Studying on Academic Performance. *The B.E. Journal of Economic Analysis & Policy*, 8(1).

Söderström, M. (2006). School choice and student achievement – new evidence on open-enrolment. *IFAU Working Paper*, 2016:16.

Söderström, M. and Uusitalo, R. (2010). School Choice and Segregation: Evidence from an Admission Reform. *The Scandinavian Journal of Economics*, 112(1):55–76.

II. Teacher credentials, per-pupil spending, and outcomes of adult education students

Acknowledgments: I am grateful to Peter Fredriksson and Helena Holmlund for their extensive comments and guidance. I also appreciate the many helpful discussions I had with Cristina Bratu, as well as the feedback that I received from Jan Sauermann, Alexander Willèn, Björn Öckert, the Uppsala Labor Group, and seminar participants at the Department of Economics at Uppsala University. Rasmus Evaldsson at Statistics Sweden deserves special thanks for the extra effort he spent locating and scanning the expenditure data from the SCB archive. Any errors in data digitization are my own.

1. Introduction

Policymakers consider increased access to education a crucial tool for improving people's skills, productivity, and well-being. However, an important question is whether access to education can expand without lowering the quality of schooling. The answer depends in part on the relationship between school resources and student achievement. Large increases in student enrollment put a strain on educational inputs such as class size, teacher quality, and per-pupil spending. For example, to meet the increased demand for teachers, schools may be forced to hire candidates with poor qualifications, thereby lowering the average quality of the teaching staff. If negative shocks to the quality or quantity of school inputs matter for student outcomes, policies that increase access to education may be limited in their effectiveness and could even have adverse consequences, particularly for students who would have enrolled in the absence of the intervention.

In this paper, I evaluate how expansion-induced shocks to school inputs affect student performance by investigating the spillover effects of a Swedish policy that temporarily doubled enrollment in adult education. The policy, known as the Adult Education Initiative (AEI), was part of a strategy to reduce high unemployment after a severe economic crisis in the early nineties. Between 1997 and 2002, the government created an additional 100,000 spots in adult education and used generous study allowances to encourage low-educated, unemployed individuals to enroll. A key feature of these study allowances is that they were available only to individuals aged 25 to 55. I exploit this institutional detail by restricting my analysis to individuals under age 25, thus mitigating concerns that changes in student composition drive my findings.

With rich administrative data covering all students in adult education and their teachers, I perform two complementary analyses. First, I study how the Adult Education Initiative affected students' exposure to a broad range of inputs that have been shown to influence academic achievement in other settings, including class size (Fredriksson et al., 2013; Krueger and Whitmore, 2001); teacher experience (Papay and Kraft, 2015); teacher certification (Andersson et al., 2011); and per-pupil expenditure (Jackson et al., 2020). Next, I evaluate whether the changes in school inputs coincided with changes in students' likelihood to complete their courses or earn high grades. Taken together, the two sets of results provide reduced-form evidence on the relationship between school inputs and student performance in adult education.

To estimate the effects of the reform, I rely on the fact that the expansion of adult education was not geographically uniform. For each municipality, I measure the degree of expansion induced by the AEI as the per-capita increase in enrollment among 25- to 55-year-olds. Then, I classify the municipalities as either a higher-expansion or lower-expansion region depending on whether they experienced above- or below-median enrollment shocks. Intuitively, my approach compares the evolution of school inputs and student outcomes in municipalities where enrollment in adult education expanded a lot (i.e., the higher-expansion regions) and municipalities where enrollment in adult education expanded a little (i.e., the lower-expansion regions). This difference-in-differences strategy is built on the idea that the amount of strain that the AEI put on school inputs should vary with the intensity of enrollment expansion.[1] Under the premise that larger increases in enrollment coincide with stronger negative shocks to school inputs, it is possible to deduce the relationship between school inputs and student outcomes by studying how academic performance evolves over time in the higher- and lower-expansion regions. If negative shocks to school inputs have a negative effect on students, then academic performance should decline in the higher-expansion regions relative to the lower-expansion regions after the introduction of the AEI.

My first set of results confirms the premise that regions subject to larger enrollment increases experienced greater strains on school inputs. Although the central government provided subsidies to help municipalities finance the expansion, the additional funding was stretched thin in areas where enrollment rose the most: I find that average per-pupil spending on instruction, course materials, and facilities all declined in the higher-expansion regions relative to the lower-expansion regions as a result of the policy. The higher-expansion regions also had a more difficult time recruiting qualified teachers. While there were large declines in the average quality of the teaching staff across both groups, my estimates show that students in the higher-expansion regions experienced a significantly larger drop—about five percentage points—in the share of teachers with a formal pedagogical background or prior teaching experience. However, there were no differential changes in class size.

My second set of results shows that, as a consequence of the AEI, students in the higher-expansion regions became four percentage points more likely to drop their courses compared to students in the lower-expansion regions. There was also a negative effect on the likelihood of earning a high grade. However, this latter effect appears to be driven by course dropout: conditional on course completion, I find no impact on grades. Together with the first set of results,

[1] For example, areas that experience larger increases in enrollment have a greater need for teachers. Given that teachers are in short supply, schools in these regions may have to crowd more students into the same classroom or hire a larger share of unqualified, inexperienced teachers from outside the profession.

these findings suggest that there is a causal link between school inputs and course completion. To support this interpretation, I also study the dynamics of the effects over time, showing that the shocks to school inputs and the changes in student outcomes follow relatively similar patterns.

While my findings are highly suggestive that school resources affect the academic achievement of adult learners, I consider several alternative explanations. Of particular concern is the fact that the composition of the student body is bound to change as a result of the reform, and these compositional changes may be larger in higher-expansion regions. This gives rise to two separate issues. First, if the AEI created more opportunities for younger students to participate in adult education in the higher-expansion regions, they may be more negatively selected than students in the lower-expansion regions, which could in turn explain the observed increase in course dropout. Indeed, I find that the higher-expansion regions experienced larger and more sustained increases in enrollment among all age groups. However, I perform a set of balance tests and show that the expansion did not have a differential effect on the composition of younger students in the higher- versus lower-expansion regions. Although I cannot rule out that some other unobserved characteristic changed in a way that would negatively impact student achievement, my balance tests lend credibility to that assumption and alleviate concerns about negative selection.

Second, in light of evidence that peer composition affects the academic achievement of college students (Carrell et al., 2009), it is plausible that peer quality could influence the performance of other types of adult learners. In that case, any differential changes in student composition among the target population—i.e., the older students—could explain my findings, especially given that the reform encouraged low-educated, unemployed individuals to enroll. Reassuringly, however, I find that, on average, the AEI affected the composition of classmates in a similar manner in the high- and low-expansion areas. Moreover, despite a few significant changes in some years, the pattern of changes in peer quality over time does not mirror the dynamics of the effects on student outcomes as closely as in the case of school resources, thus supporting the interpretation that shocks to school resources and not peer characteristics cause the observed increase in course dropout.

The findings of this study contribute to a broad literature on school inputs and student outcomes. Most existing studies look at the effect of resources in primary and secondary school. There are a few studies at the college level (see e.g. Hoffmann and Oreopoulos, 2009), but to the best of my knowledge, there has been no prior research on adult education students. This is an important omission, given that between 5 and 15% of the adult population in OECD countries participates in formal adult education (see p. 316 of OECD, 2017), and many of these adult learners are vulnerable members of society, for example refugees and high school dropouts, who may be unable to compensate for poor school environments.

Another key difference between my study and the existing research is that most other studies try to isolate the effect of one particular school input when holding other school inputs constant. While this allows for cleaner causal identification, it does not reflect the reality of most educational interventions, i.e., that many inputs may change at once. One notable exception in the literature is Jepsen and Rivkin (2009), who study a large-scale class-size reduction in California and show that the benefits of smaller classes can be offset when schools have to hire inexperienced, uncertified teachers in order to meet class size targets. My findings echo these results, suggesting that the benefits of educational expansions may be diminished by resulting shocks to school resources.

As such, my study also contributes to the literature on educational expansions, providing some of the first quasi-experimental evidence that educational expansions have negative spillover effects. In a closely related study, Bianchi (2020) evaluates an expansion of undergraduate STEM education in Italy and finds negative effects on the academic performance of students who were not the target of the policy. Similar results have been found in the literature on cohort size and resource crowding (see, e.g., Babcock et al., 2012; Bound and Turner, 2007). However, as far as I know, none of the existing studies look specifically at enrollment expansions in the adult education sector. This is a topical issue, as enrollment in adult education is on the rise in most OECD countries and policymakers have acknowledged that “lifelong learning” is a key policy tool to cope with technological changes on the labor market.

The rest of the paper proceeds as follows. The next section provides an overview of the Swedish education system and the Adult Education Initiative, followed by a discussion of the data and the key variables used in my analysis. Section 3 discusses the empirical strategy, sample selection, and descriptive statistics. Section 4 reports the results of the difference-in-differences analysis for school inputs and the corresponding results for students’ academic outcomes. Section 5 presents some robustness checks, and section 6 performs heterogeneity analysis by individual’s background characteristics. Finally, section 7 concludes.

2. Context and data

All of the facts presented in this section apply specifically to my period of study (1994-2002). I begin with an overview of the education system in Sweden, in particular municipal adult education (*Komvux* in Swedish). Then, I proceed with a discussion of the Adult Education Initiative, the policy intervention that I exploit to study the relationship between school inputs and student outcomes. Finally, I present the data sources and key variables that I use for my empirical analysis.

2.1 The Swedish education system

Following nine years of compulsory education, the majority of students in Sweden choose to enroll in high school. High school education is divided into specialized tracks that are either academic or vocational in nature. Until the mid-1990s, vocational tracks lasted two years and did not grant eligibility for university admission, whereas academic tracks lasted three years and prepared students for higher education. By 1996, the vocational tracks had been converted to three-year programs, and all high school graduates met the general admission requirements for university. However, some students had to complete additional courses in order to become eligible for university programs with special entry criteria.[2]

Once individuals complete high school or reach age 20, they are eligible to enroll in municipal adult education in their municipality of residence. They can request to enroll in other municipalities under special circumstances, for example if their home municipality does not offer certain courses. At the lower-secondary level, admittance is guaranteed to any student who has not finished compulsory school. At the upper-secondary level, admittance is guaranteed only when there is sufficient capacity in a course. If demand for a course exceeds the number of available spots, the school chooses which applicants to admit according to national guidelines: priority is given to those who lack a three-year high school degree and to those in greatest need of studying the course.[3] If a student is admitted, the municipalities must provide the education free of charge. Moreover, the central government offers various forms of financial aid to help students cover their living expenses and foregone earnings while enrolled.

Enrollment in adult education is quite common in Sweden, with over a third of a birth cohort participating at some point in their life. There are several

[2]For example, medical programs require specific courses in math and science.

[3]The Ordinance on Municipal Adult Education (*Förordning om kommunal vuxenutbildning, SFS nr. 1992:403*) outlines the admission guidelines in more detail.

common reasons for participating. Compulsory or high school dropouts may enroll to complete their degree, or graduates of two-year vocational tracks may register for the additional courses required to top up to a three-year degree. Individuals with ambitions to attend a particular university program may enroll to complete courses that were not part of their high school track but are required for admission to their desired field of study. During the period I study, it was also possible for high school graduates to sign up for courses they had already completed in high school in an attempt to improve their final grade and boost their chances of college admission. Lastly, students who want additional occupational training may enroll in specialized vocational courses to supplement their previous training.[4]

The vast majority of enrollment in adult education—around 85%—occurs in upper-secondary courses, with only 10% in the compulsory-school level and just 5% in supplementary vocational training. Almost all courses follow a syllabus that is similar to—or in the case of upper-secondary courses, identical to—the syllabus in the regular school system. The National Agency for Education (*Skolverket*) determines both the syllabus and the grading criteria. At the compulsory level and in supplementary vocational training, teachers can assign three grades—Fail (*I*), Pass (*G*), and Pass with Distinction (*VG*)—while at the upper-secondary level, teachers can also pass a student with High Distinction (*MVG*). If teachers lack sufficient basis to judge a student's mastery of the subject (e.g., due insufficient course participation), the teacher is not supposed to set a formal grade, but should instead enter a mark of *Z* into the grading catalogue.

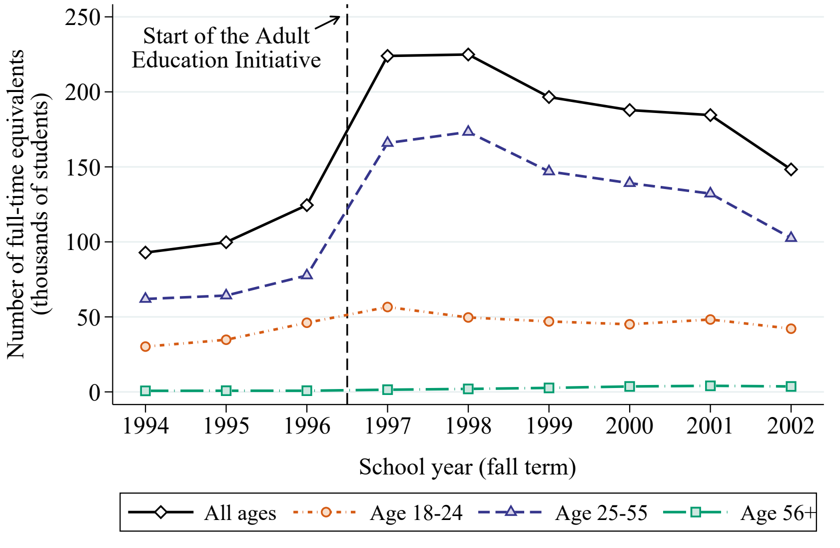
2.2 The Adult Education Initiative

Between 1997 and 2002, the adult education sector underwent a massive expansion as a result of an intervention called the Adult Education Initiative (AEI). The Swedish government implemented the policy in response to a severe financial crisis that caused unemployment to skyrocket from under 2% in 1990 to over 8% percent by the mid-1990s. The primary aim was to reduce unemployment among the low-educated by giving them a chance to obtain stronger academic credentials and raise their appeal to potential employers. In addition, the initiative was intended to revitalize the provision of adult education.

[4]These supplementary courses are called *påbyggnadsutbildningar* in Swedish. If the course is a continuation of specific training received in high school or another course in adult education, national guidelines stipulate that grades should be used for admission.

To achieve its goals, the government financed the creation of 100,000 spots in municipal adult education, primarily at the upper-secondary level. Within just two years of the program's start in July 1997, enrollment in adult education nearly doubled. Figure 1 shows that much of this increase resulted from a sharp jump in enrollment among individuals between ages 25 and 55. The government specifically targeted this age group with generous study allowances: for up to one year, low-educated 25- to 55 year olds who were eligible for unemployment benefits could instead receive the same amount in study aid.[5]

Figure 1. Level of enrollment in adult education over time by age group.



Notes: Enrollment levels are measured in full-time equivalent (FTE) students. I follow Statistics Sweden's definition and calculate FTEs as the total number of registered lecture hours divided by 540. Age is measured on December 31st of the reported year.

Although the central government was in charge of financing the initiative, both in terms of providing financial aid to the students and subsidies to the municipalities, the municipal government bore the ultimate responsibility for implementing the policy. Municipalities had a large degree of freedom in determining which organizational committee would oversee the reform;[6] the

[5]These study allowances were called UBS. See p. 46 of Stenberg (2010) for more details on the eligibility requirements.
[6]Some municipalities created special committees to carry out the administrative oversight, while others relied on the principals already in charge of organizing Komvux.

number and type of courses that they would offer;[7] and the extent to which they would hire external providers to assist with course instruction. However, all municipalities were subject to several key requirements. First, in order to receive government subsidies, they had to maintain the same "base organization" (i.e., enrollment level) that they had in the years leading up to the expansion. Second, they had to—at least in principle—follow separate ordinances for admitting AEI students and regular Komvux students.[8] In practice, however, these separate admission procedures appear to have been difficult to follow, and there was some arbitrariness in whether students were officially counted as AEI participants or part of the base organization.[9] This makes it difficult to assess the exact extent to which younger students and AEI participants enrolled in the same classes. To provide some idea, Figure A.1 shows the age composition of classes prior to and during the AEI. While there appear to be some AEI-exclusive classes, the figure suggests that classes remained relatively mixed after the expansion. This has two implications for my analysis. On the one hand, it highlights the importance of checking for changes in peer composition, which is a competing explanation to the hypothesis that shocks to school inputs drive my findings. On the other hand, if the classes had not been integrated, it would have been easier for schools to target their resources, e.g., more qualified teachers or funding, at specific students. With integrated courses, the average shocks that I estimate for municipal- and school-level inputs are more likely to capture the actual shock to school inputs faced by younger students.[10]

2.3 Data sources and definition of key variables

This paper uses administrative data from Sweden covering all participants in municipal adult education between 1994 and 2002. Several key variables come from the Komvux registry (*Komvuxdatabasen*), which contains detailed enrollment history and course transcripts for the full population of students. For each course that a student enrolls in, I observe the total number of lecture

[7] The initiative aimed to promote the accumulation of general skills rather than vocation-specific skills, but the government encouraged municipalities to adjust course offerings based on the demands and needs of their residents.

[8] Similar to the rules described earlier, there was a specific order for admitting students to oversubscribed courses: the key difference was that, in the case of AEI students, preference was given to *unemployed* individuals who lacked a three-year high school degree.

[9] An explanation is provided on p. 38–39 of Skolverket's first official annual evaluation of the AEI (Skolverket, 1998). See also Gotlands kommun (2001) for anecdotal evidence.

[10] Due to data limitations, I cannot link teachers directly to their students, nor can I see how much money each school spent per student in municipalities with more than one provider of adult education.

hours, an indicator of whether the individual unregistered from the course, and their final grade. The variable for lecture hours allows me to calculate the number of full-time student equivalents and capture the variation in enrollment over time and across regions. This regional and temporal variation in enrollment is key for my identification strategy. The variables for course grades and de-registration provide measures of academic performance that serve as the main student outcomes in my analysis. All of these student outcomes are defined at the course level and are binary variables equal to one if a condition is met and zero otherwise. For example, course dropout equals one if the individual does not earn a grade in the course[11] and zero if they obtain any grade. I also create indicators for receiving any passing grade in the course and for receiving honors.[12] I do not have data on course credits prior to 1997, so I am unable to study credit accumulation.

The Komvux registry includes a school code that allows me to link students to the adult education teachers employed at their school at the start of the academic year. Via the national teacher registry (*Lärarregistret*), I obtain annual information on teachers' certification status and accumulated years of teaching experience since 1985. I also extract information on teachers' completed years of schooling via the Integrated Database for Labor Market Research (*LOUISE*). I use the data on these three teacher characteristics to measure the average teacher qualifications that students are exposed to. I do not observe the exact courses taught by each teacher, so I construct school-level averages of the characteristics, weighing each teachers' characteristic by their percent of employment such that more weight is given to the qualifications of full-time teachers than part-time teachers.[13]

In addition to the school code, the Komvux registry also includes course codes and detailed course information that enable me to approximate "classes" of students who study a course together (see the data appendix for details). I use this information to measure class-level peer quality and class size, which I define as the number of registrants at the start of the course.[14] Unfortunately, I do not have data on prior academic achievement for the majority of adult education students during the period I study, because Statistic Sweden's

[11] For example, if they unregister, or if they fail to participate enough for the teacher to assign them a grade.

[12] Upper-secondary courses have two honors grades—Distinction and High Distinction—whereas the others only have one honors grade. Thus, for upper-secondary courses, I consider both Pass with Distinction and High Distinction as receiving honors.

[13] The teacher registry contains all employees with valid contracts in October. Thus, if schools hire new teachers during the spring term, these teachers are excluded from the averages.

[14] When defining class size, I do not exclude students who de-register from the course at some point during the semester, because this behavior is likely endogenous to the reform.

compulsory school and high school registries only date back to the late 1980s. However, I can proxy individuals' ability using several different measures: own education level, parents' education level, percentile rank in the national wage distribution for their birth cohort, and days of unemployment during the year prior to enrollment. I also study the average age of peers and the share of female peers in a class.

As a complement to the non-financial school inputs that I study, I also collect data on each municipality's expenditure on adult education. The data for 1994 through 1998 came in paper form from Statistics Sweden's archive and had to be digitized, whereas the National Agency for Education delivered the data in digital form for school years 1999 through 2002. All variables are reported at the municipal level and are measured as costs per full-time-equivalent students. My analysis looks at the log of per-pupil expenditure on instruction, learning materials, and learning facilities.

3. Empirical strategy

This paper exploits enrollment shocks induced by the Adult Education Initiative (AEI) to generate plausibly exogenous variation in school inputs. The crux of the identification strategy is that regions subject to larger increases in enrollment as a result of the AEI experience stronger negative shocks to school inputs. Under this premise, it is possible to assess the impact of school inputs on student outcomes by studying how student performance evolves over time in regions subject to higher versus lower enrollment shocks. If school inputs matter, student outcomes should decline in higher-expansion regions relative to lower-expansion regions after the introduction of the AEI.

A potential issue with this empirical strategy is that the educational expansion I exploit is likely to change the composition of students enrolled in adult education. It is reasonable to expect that the average ability level declines with the influx of new students, and these declines are likely to be stronger in areas where enrollment expands the most. Any observed changes in student performance might therefore reflect changes in students' own underlying academic ability or the quality of their peers.

One crucial feature of the AEI allows me to address concerns related to negative selection. As discussed earlier, the intervention primarily targeted low-educated, unemployed individuals aged 25 to 55 by incentivizing their enrollment with generous study allowances. By contrast, there were no significant changes in the financial incentives or admission rules for younger students. This means that selection issues are likely a big concern among the older population, but less so for the younger students. I therefore restrict my analysis to individuals under age 25 at the time of their initial enrollment, and in section 3.3, I provide a formal test to show that—at least on observed characteristics—there are no significant compositional changes in the higher- relative to the lower-expansion regions among this subsample. Moreover, as a robustness check in section 5, I will study how the composition of classmates changes due to the AEI and show that the differences in compositional changes between the higher- and lower-expansion regions are surprisingly minor and do not appear to drive my findings.

3.1 Identifying variation

In order to implement my identification strategy, I need to define a measure that captures regional variation in the intensity of enrollment shock caused by the Adult Education Initiative. Since the AEI targeted individuals aged 25 through 55, I focus on the variation in enrollment for this age group. I measure the intensity of expansion in each municipality as the difference between average per-capita enrollment amongst 25-55 year olds during the eleven academic terms that the AEI was in place (fall term 1997 through fall term 2002) and the six academic terms prior to the AEI (fall term 1994 through spring term 1997). That is, the enrollment shock for a given municipality m is defined as:

$$Expansion_m = \sum_{t=0}^{10} \frac{1}{11} \cdot \frac{Enrollment25to55_{m,t}}{Population25to55_{m,t}} - \sum_{t=-6}^{-1} \frac{1}{6} \cdot \frac{Enrollment25to55_{m,t}}{Population25to55_{m,t}} \tag{1}$$

where subscript t indexes the academic term relative to the start of the Adult Education Initiative (i.e., $t = 0$ in fall 1997); the variable $Enrollment25to55_{m,t}$ is the number of full-time-equivalent students aged 25 through 55 registered in municipality m during term t; and the variable $Population25to55_{m,t}$ is the number of individuals aged 25 to 55 residing in municipality m (measured in hundreds). The higher the value of $Expansion_m$, the larger the enrollment shock.

Panel (a) of Figure A.2 shows the variation in enrollment shocks across different municipalities. For ease of exposition, I divide the municipalities into two groups in the remainder of my empirical analysis. The higher-expansion group consists of the 143 municipalities that experienced above-median enrollment shocks, and the lower-expansion group consists of the 143 municipalities that experienced below-median enrollment shocks. Panel (b) of the figure shows the municipalities according to this binary classification.

An essential question is whether the higher-expansion municipalities experience sufficiently large enrollment shocks relative to the lower-expansion municipalities, such that we should expect the strain on school resources to be larger in the higher-expansion regions. I will analyze this more formally in the results section, but to convey the intuition, Figure 2 plots the strength of the enrollment shocks in the two groups and provides an example to illustrate that greater increases in enrollment indeed coincide with stronger negative shocks to school inputs. In the top panel, we see that prior to the start of the AEI, enrollment per capita in the higher- and lower-expansion regions was essentially equal. When the AEI began in 1997, the enrollment level jumped in both regions, but the shock was much larger in magnitude in the higher-expansion regions. These differences persisted through the end of the AEI, and even widened slightly between 1999 and 2001 as enrollment tapered off more quickly in the lower shock regions. In the bottom panel, we see a similar pattern in students' exposure to qualified teachers. Prior to the AEI, students in the higher- and lower-expansion regions were taught by teachers with similar credentials: on average, about 85% of teachers in their school were certified. This percentage dropped sharply in both regions as the increased demand for teachers meant that municipalities had to hire teachers without a pedagogical background (see Figure A.3 in the appendix). However, the declines in teacher qualifications were steeper in the higher-expansion region, particularly during the 1999-2001 school years, when enrollment levels declined relatively faster in the lower-expansion regions.

3.2 Difference-in-differences specification

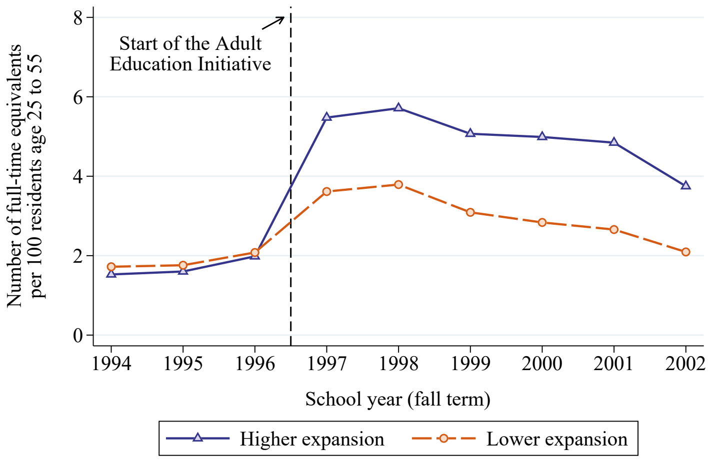
The preceding example suggests that the binary classification of municipalities by above-median and below-median enrollment expansion accurately captures school resource shocks. To more formally compare the evolution of outcomes in higher- and lower-expansion regions, I use the following difference-in-differences model:

$$\begin{aligned} Outcome_{i,c,s,m,y} = {} & \gamma(HigherExpansion_m \times PostAEI_y) \; + X'_{i,y}\psi \\ & + \alpha_m + \beta_y + \omega_s + \varepsilon_{i,c,s,m,y} \end{aligned} \tag{2}$$

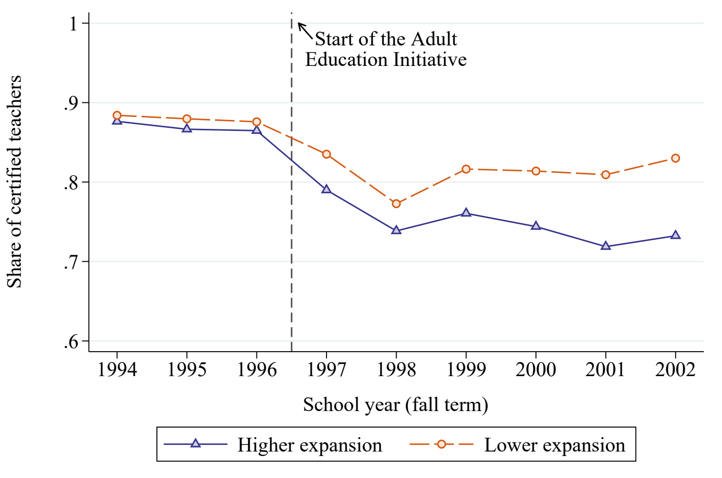
where subscript i indexes an individual; c indexes a course; s indexes a subject; m indexes the municipality of enrollment; and y indexes the school year. In the first part of my analysis, the dependent variable $Outcome_{i,c,s,m,y}$ measures student i's exposure to various school inputs, and in the second part, it

Figure 2. Illustration of the identification strategy.

(a) Enrollment shocks per capita



(b) Students' exposure to certified teachers



Notes: Higher-expansion regions experienced above-median enrollment shocks, and lower-expansion regions experienced below-median enrollment shocks. In panel (b), certified teachers are those with a college degree in pedagogy.

measures student i's achievement in course c. The indicator $PostAEI_y$ equals one for all school years after the introduction of the Adult Education Initiative (1997-2002), and $HigherExpansion_m$ equals one for municipalities that experienced an above-median intensity enrollment shock as a result of the policy. For all outcomes, the specification includes municipality fixed effects (α_m) and school year fixed effects (β_y). Because I pool observations for all courses, I also include subject fixed effects (ω_s) in the regressions for student outcomes.[15] These fixed effects account for unobserved heterogeneity due to the fact that some subjects may be inherently harder to pass than others, e.g., if subjects like mathematics are systematically graded more harshly than subjects like arts.

The parameter of interest, γ, measures how outcomes evolved in the higher-expansion regions relative to the lower-expansion regions after the introduction of the AEI. It captures the average effect of the AEI on school resources and student outcomes under the assumption that the outcomes would have followed parallel paths in the absence of the intervention. To explore the pattern of the effects over time, I also estimate the following event study specification where I replace *PostAEI* in equation (2) with a set of time dummies:

$$\begin{aligned} Outcome_{i,c,s,m,y} = & \sum_{y=1994}^{1995} \lambda_y \cdot HigherExpansion_m + \sum_{y=1997}^{2002} \lambda_y \cdot HigherExpansion_m \\ & + X'_{i,y}\psi + \alpha_m + \omega_s + \varepsilon_{i,c,s,m,y} \end{aligned} \tag{3}$$

The coefficients of interest, λ_y, are normalized with respect to the year prior to the AEI. In addition to shedding light on the dynamics of the effects, this specification allows me to evaluate whether the parallel trends assumption is credible: if so, the pre-AEI coefficients λ_{1994} and λ_{1995} should be statistically indistinguishable from zero.

3.3 Sample selection and descriptive statistics

To construct my sample, I start with the full population of students enrolled in municipal adult education between fall term 1994 and fall term 2002. As discussed earlier, I first restrict my analysis to individuals who were age 18 to 24 at the time of initial enrollment in order to mitigate concerns that changes in student composition drive my results. Additionally, I restrict the sample

[15] I define subjects using a combination of the prefix in the course code (e.g., MA for mathematics) and the level of study (e.g., lower secondary or upper secondary).

to individuals with valid compulsory school grades, so that I can control for prior achievement levels in the regressions to further alleviate concerns that differences in underlying ability affect my findings. If an individual enrolls in the same course multiple times, I keep the observation corresponding to the earliest enrollment. Lastly, I drop schools where I cannot link to teacher characteristics and classes with unreported grades.[16] The resulting sample consists of 314,408 individuals, and 1,989,322 observations at the course level.

Tables A.1 and A.2 in the appendix show the descriptive statistics for the sample of students in higher- and lower-expansion regions. The two groups are relatively similar, though students in the higher-expansion regions are slightly less likely to have a foreign background and have slightly weaker performance in both compulsory school and high school. While it is not essential that students' average characteristics are identical in the two groups, an underlying assumption of my identification strategy is that group composition does not change differently across higher- and lower-expansion regions in a way that is correlated with student outcomes. Of particular concern is the fact that higher-expansion regions experienced larger and more sustained increases in enrollment among 18-to 24-year olds, not just the target population (see Figure A.4).

Before proceeding to the main empirical analysis, I provide evidence that in spite of these different enrollment patterns, there were no differential changes in the composition of students in the higher-expansion and lower-expansion regions. To this end, I estimate the difference-in-differences model in equation (2) using students' background characteristics as the dependent variable. Table 1 reports the results of the balance tests. The test for the joint significance of all characteristics has a p-value of 0.643, indicating no significant changes between the two groups. This is confirmed by the separate tests for the significance of each coefficient. All point estimates are small in magnitude, and every estimate is statistically indistinguishable from zero at conventional significance levels, with the exception of the estimate for grade nine GPA at the 10% level. Given the number of variables that I test, this could be due to random chance; indeed, the more recent achievement measures like high school completion and high school GPA suggest no differential changes in academic

[16] When a student is still registered for a course but does not attend enough lectures or turn in assignments for a final grade, teachers are supposed to record a grade of Z rather than a missing value. However, grades are missing for approximately 10-20% of the observations each year. Some of these are for valid reasons, e.g., introductory courses where students are never assigned grades. However, for most courses, it is impossible to know whether the grades "should" be missing or whether the teacher failed to record a grade (Z or otherwise). To be conservative, I drop all classes with unreported grades, but the main findings are unchanged when I relax this restriction. See the data appendix for additional details.

Table 1. Balance of covariates.

	Estimate	Std. Err.	P-value
Female	0.005	0.006	0.413
Age	0.034	0.041	0.418
Married	-0.001	0.002	0.707
Born in Sweden	-0.002	0.003	0.564
Swedish-born mother	0.002	0.005	0.652
Swedish-born father	0.002	0.004	0.683
Mother's years of schooling	-0.028	0.031	0.370
Father's years of schooling	-0.013	0.045	0.765
Grade 9 GPA	-0.025	0.015	0.093
High school GPA	-0.014	0.013	0.295
Graduate of high school	0.004	0.007	0.525
Graduate of academic program	0.001	0.009	0.908
Graduate of vocational program	0.005	0.008	0.536
Graduate of special program	-0.002	0.002	0.429
P-value for test of joint significance:			0.643

Notes: Each entry of the second column reports the estimate of the interaction term in equation (2) from separate regressions where the corresponding characteristic in the first column is the outcome variable. Grade nine and high school GPA are standardized by graduation cohort. Standard errors in the third column are clustered at the municipal level.

ability between the two groups over time. Nevertheless, I include grade nine GPA and other control variables in all of my main regressions to control for any differences.

Before proceeding to the analysis, it is also relevant to investigate whether higher- and lower-expansion municipalities have similar background characteristics. My model specification in equation (2) includes municipality fixed effects to control for time invariant differences across municipalities that might be related to both the level of adult education expansion and student outcomes. Thus, it is not necessary for the enrollment shocks I exploit for identification to be unrelated to municipal characteristics. Nevertheless, significant differences in municipal characteristics could suggest that the trends in student outcomes might have diverged even in the absence of the Adult Education Initiative, which would violate the parallel trends assumption.

The map in Figure A.2 shows that the municipalities in the two groups are reasonably spread out across Sweden, though the lower-expansion regions are slightly more concentrated in metropolitan areas. I report the characteristics of these municipalities in Table A.3. The industry structure in the two regions is relatively similar, but residents of higher-expansion regions face slightly worse labor market conditions and are slightly less educated on average. To deal with concerns that these differences might reflect different underlying trends in student outcomes, I will perform robustness checks where I control for trends in municipal characteristics over time.

4. Difference-in-differences analysis

4.1 Effects on school inputs

My empirical strategy rests on the premise that regions subject to larger enrollment increases following the introduction of the Adult Education Initiative experience stronger negative shocks to school inputs. The descriptive plot in Figure 2 suggests that this is the case, at least for teacher certification status. In this section, I provide more formal evidence by estimating the difference-in-differences model in equation (2) using various school inputs as the dependent variable. I analyze several types of inputs: class size, teacher credentials, and log per-pupil expenditure.

Table 2 presents the results of the difference-in-differences analysis. The first row reports the estimate of the interaction term, γ, which captures how school inputs evolved in regions with higher expansion intensity relative to regions with lower expansion intensity after the start of the AEI. The results confirm that the stronger the enrollment shock, the stronger the negative shock to school inputs. Relative to students in regions that experienced lower enrollment shocks, students in the higher-expansion regions were taught by teachers who were less educated and more likely to be inexperienced and uncertified. The higher-expansion municipalities also spent less money per student on the cost of instruction. Although the effects are insignificant, there also appear to be declines in per student expenditure on learning materials and school facilities. Class size, however, remained unchanged as a result of the reform.

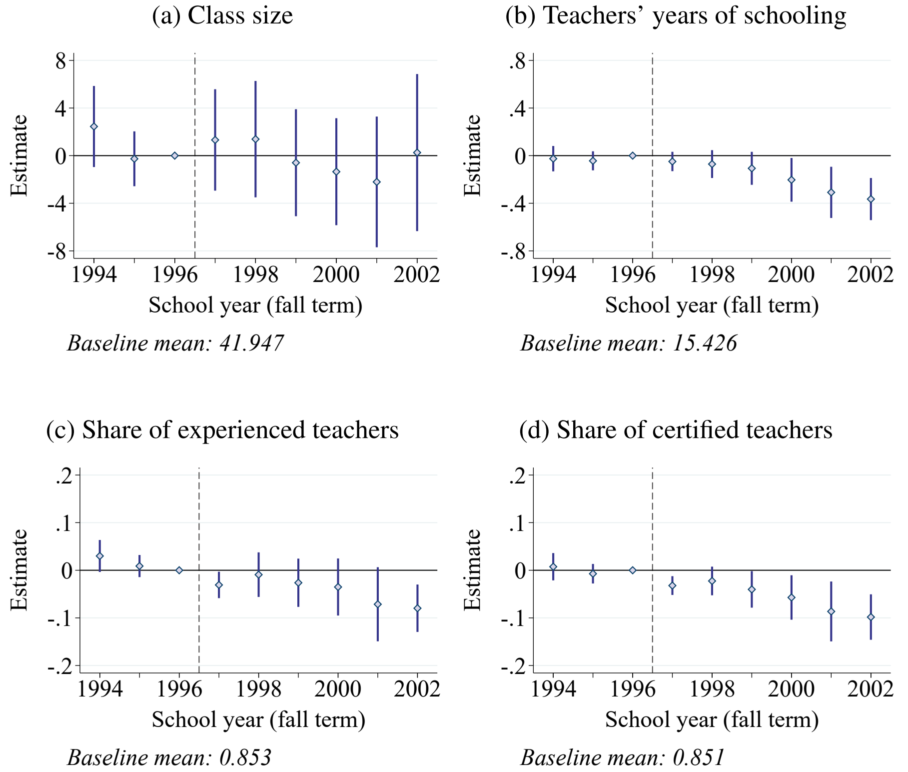
The event study plots in Figure 3 and Figure 4 confirm these findings and shed light on how the inputs evolve pre- and post-reform. The figures report the annual coefficients from the estimation of equation (3) for each school input. All of the pre-reform coefficients are statistically indistinguishable from zero, suggesting that the inputs would have evolved similarly in the absence of the enrollment shock caused by the Adult Education Initiative. The higher-expansion regions immediately experience a stronger strain on school resources after the implementation of the AEI, but these shocks grow stronger over time for the majority of characteristics. If school inputs matter for student achievement, we should expect to see similar dynamics in the effects for student outcomes.

Table 2. Effects on school inputs.

		Teacher credentials			Log per-pupil expenditure		
	Class size (1)	Years of Schooling (2)	Share experienced (3)	Share certified (4)	Cost of instruction (5)	Learning materials (6)	School facilities (7)
DD estimate	-0.548	-0.145	-0.051	-0.048	-0.111	-0.140	-0.211
	(2.034)	(0.051)***	(0.014)***	(0.019)**	(0.047)**	(0.104)	(0.131)
Pre-reform mean	40.644	15.445	0.855	0.849	9.702	7.309	8.431
Observations	1,989,351	1,989,351	1,989,351	1,989,351	1,947,775	1,940,322	1,941,854

Notes: Each individual appears once per course. Log per-pupil expenditure is calculated at the municipal level and measured in terms of full-time equivalents. The teacher characteristics are calculated at the school level. Experienced teachers are defined as teachers with 3+ years of teaching experience, and certified teachers are teachers with a college-level degree in pedagogy. The number of observations for the per-pupil expenditure outcomes differs from the other columns due the fact that some municipalities did not report expenditure data to the National Agency for Education. The pre-reform mean refers to the average in the higher-expansion regions during the 1994 to 1996 school years. Standard errors are shown in parentheses and clustered at the municipal level. Stars denote significance levels: *** for $p < 0.01$; ** for $p < 0.05$; * for $p < 0.10$.

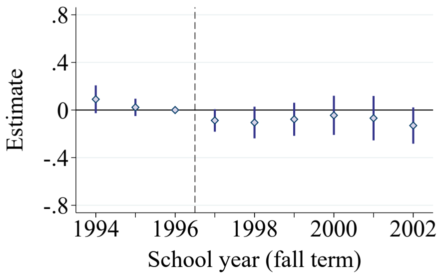
Figure 3. Effects on school inputs over time.



Notes: In all panels, the vertical line indicates the introduction of the Adult Education Initiative, and the baseline mean refers to the average in the higher-expansion regions in 1996. Each point plots the estimates of λ_y from equation (3), and the blue bars show the 95% confidence intervals for the estimates when clustering the standard errors at the municipal level. In panel (a), class size refers to the number of students registered in a class at the start of the term. In panels (b) through (d), the teacher characteristics are calculated at the school level. Experienced teachers have at least three years of experience. Certified teachers are those who hold a pedagogy degree.

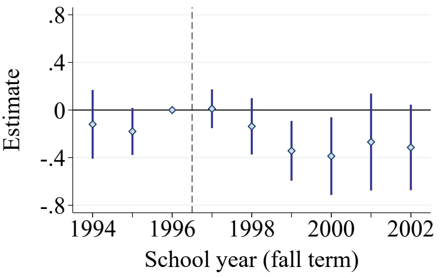
Figure 4. Effects on log per-pupil expenditure over time.

(a) Instruction



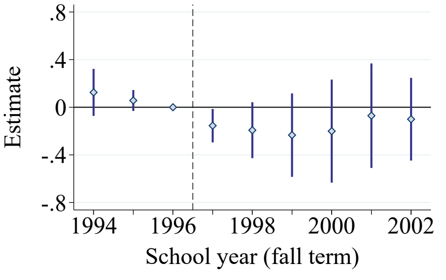
Baseline mean: 9.662

(b) Learning materials



Baseline mean: 7.414

(c) Facilities



Baseline mean: 8.399

Notes: In all panels, the vertical line indicates the introduction of the Adult Education Initiative, and the baseline mean refers to the average in the higher-expansion regions in 1996. Each point plots the estimates of λ_y from equation (3), and the blue bars show the 95% confidence intervals for the estimates when clustering the standard errors at the municipal level. In the raw data, costs are reported at the municipal level in thousands of Swedish crowns per full-time-equivalent student enrolled in the municipality. I further transform each outcome by taking the natural logarithm. Standard errors are clustered at the municipal level.

4.2 Effects on course outcomes

The previous section established that students in regions where the Adult Education Initiative induced higher enrollment shocks were taught by less-qualified teachers than students in regions that experienced lower enrollment shocks. Additionally, per-pupil expenditure declined in the higher-expansion regions relative to the lower-expansion regions. If these school inputs have an impact on students' academic achievement, the outcomes of students in the higher-expansion regions should decrease relative to the outcomes of students in the lower-expansion regions after the start of the AEI. To investigate this, I repeat the difference-in-differences analysis specified in equation (2) using different course outcomes as the dependent variable. In particular, I look at the probability of dropping out of a course, receiving any passing grade in the course, or passing the course with honors.

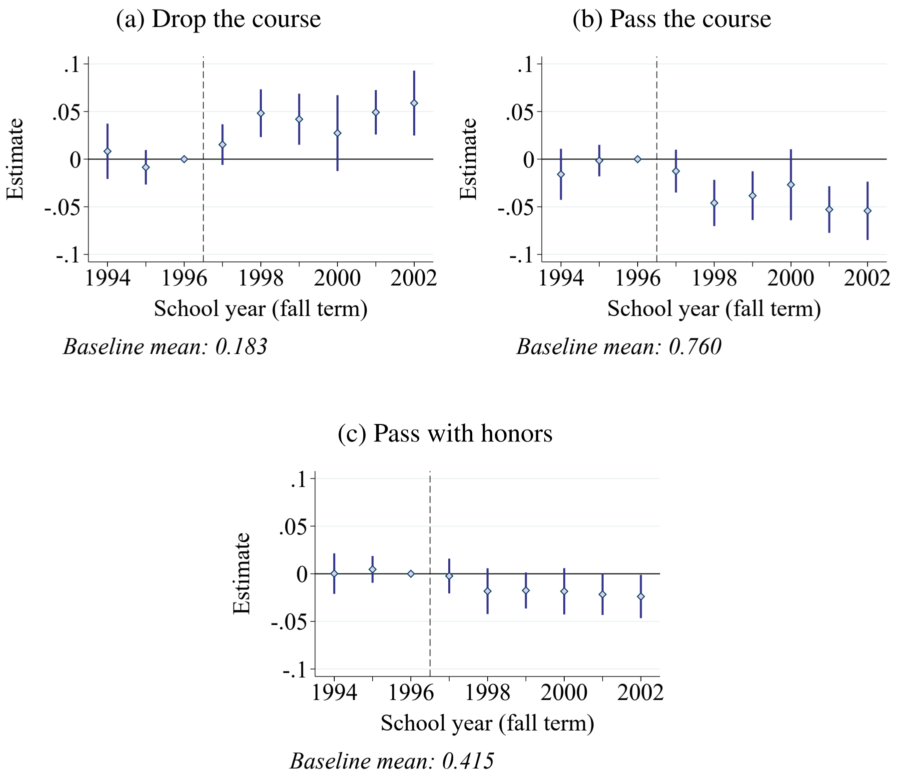
Table 3 reports the estimate of the interaction term γ for each outcome variable. Column 1 shows that as a result of the AEI, students in higher-expansion regions became four percentage points more likely to drop out of a course relative to students in lower-expansion regions. This is a sizable effect, about a 21% increase over the baseline probability of dropout. Columns 2 and 3 show that students in higher-expansion regions also became less likely to pass the course—at all or with honors—although the results in Table A.4 show these decreases are driven by the increased dropout rates rather than a negative impact on course grades.

Table 3. Effects on students' course outcomes.

	Drop the course (1)	Pass the course (2)	Pass with honors (3)
DD estimate	0.040	-0.033	-0.017
	(0.009)***	(0.009)***	(0.008)**
Pre-reform mean	0.190	0.747	0.400
Observations	1,989,322	1,989,322	1,989,322

Notes: All outcomes are unconditional probabilities. The regressions include year, municipality, and subject fixed effects, where a subject is defined as a combination of course prefix and level of study (e.g., mathematics at the upper-secondary level). They also control for all individual characteristics listed in Table A.1, except for high school GPA, which is missing for dropouts (ca. 24% of the sample). Standard errors are clustered at the municipal level. Pre-reform mean refers to the mean in the higher-expansion regions during the 1994-96 school years. Standard errors are shown in parentheses and clustered at the municipal level. Stars denote significance levels: *** for $p < 0.01$; ** for $p < 0.05$; * for $p < 0.10$.

Figure 5. Effects on course outcomes over time.



(a) Drop the course

Baseline mean: 0.183

(b) Pass the course

Baseline mean: 0.760

(c) Pass with honors

Baseline mean: 0.415

Notes: In all panels, the vertical line indicates the introduction of the Adult Education Initiative, and the baseline mean refers to the average in the higher-expansion regions in 1996. Each point plots the estimates of λ_y from equation (3), and the blue bars show the 95% confidence intervals for the estimates when clustering the standard errors at the municipal level. All outcomes are unconditional probabilities. Individuals appear once for every course that they take. Regressions include individual controls, and the following fixed effects: municipality; year; and subject. Standard errors are clustered at the municipal level.

The estimates reveal that, on average, students in higher-expansion regions had higher dropout rates than students in lower-expansion regions as a result of the AEI. This is suggestive that school inputs do have some impact on academic outcomes. To investigate this more closely, I study whether the dynamics of the effects for course outcomes line up with the patterns for school inputs that we observed in Figures 3 and 4. To this end, I estimate the event study specification in equation (3) for each of the course outcomes and plot the coefficients over time in Figure 5.

Reassuring for the identification strategy, the coefficients in the pre-reform years are again statistically indistinguishable from zero, which lends credibility to the assumption that students' course outcomes would have evolved similarly in the higher- and lower-expansion regions if they had not been subjected to the enrollment and resource shocks. After the introduction of the AEI, we see immediate declines in student performance, with the effects growing over time in a similar pattern to the pattern of effects for school inputs. The fact that the dynamics of the effects for school inputs and student outcomes follow one another so closely is highly suggestive of a causal link.

5. Sensitivity and credibility of the results

5.1 Robustness checks

According to national guidelines, teachers are not supposed to assign a formal grade but rather are supposed to report a grade of Z in the grade catalogue when students fail to attend sufficient lectures or submit assignments required to judge their mastery of the subject material. However, in the Komvux registry, grades are missing for around 15% of observations where one would expect grades to be reported. It is unclear whether these ought to be grades of Z or whether they are truly missing values. In my main analysis, I take a restrictive approach and drop all classes where any student in the class is missing a grade. As a first robustness check, Figure A.5 plots the difference-in-differences estimates when dropping courses with different shares of missing values for grades and shows that my findings are not driven by my sample restriction on missing grades.

Furthermore, I perform several specification checks to verify my main findings. Table A.5 reports the results of the different checks and includes the estimates from the main specification for comparison.

The first two robustness checks test whether my findings are sensitive to the definition of the treatment variable. In panel B, I use a continuous measure of the enrollment shock instead of the binary indicator. The effects are smaller in magnitude, though at the average treatment intensity of 2.1, they are fairly similar in magnitude to the main point estimates.

In panel C, I again use a binary treatment indicator that divides the municipalities into groups based on above-median and below-median treatment intensity; however, instead of measuring the enrollment shock as in equation (1), I take the difference between the enrollment level in the peak post-reform school year (1998/99) and the last pre-reform school year (1996/97) to better capture the immediate shock of the reform. There is a reduction in the magnitude of the effects, though they are still sizable and with the exception of the effect on passing with honors, statistically significant at the 1% level.

As an additional robustness check, I interact the set of municipality characteristics listed in Table A.3 with the time variable to allow for the possibility that there are different underlying trends in student outcomes related to municipality characteristics. The effect sizes drop considerably, but the findings point in the same direction.

In panel E, I check whether selective re-location is a problem by assigning individuals to treatment based on their municipality of residence at the start of the sample period. Re-scaling the intent-to-treat estimates by the probability of still living in the same region, the estimates are quite similar to the main specification.

5.2 Alternative explanations

While the findings presented thus far provide suggestive evidence that shocks to school inputs have a causal effect on course dropout, there are several other plausible explanations to consider. As with any large-scale educational expansion, the overall composition of the student body is bound to change due to the reform. Given that the intervention targeted older individuals without a three-year high school education, we would expect the average age of adult education students to increase and their average education level to decrease. A key question is whether these changes are stronger in the higher-expansion regions relative to the lower-expansion regions. If so, then changes in peer composition may drive my findings.

I perform two analyses to check whether the increased dropout rates are the result of changes in classmates' characteristics. First, I estimate the main difference-in-differences model using various peer characteristics as the outcome variable, and also use the event study specification to check whether the dynamic of any changes in peer composition matches the pattern of the effects that we observe for student outcomes. Second, even though they are endogenous to the reform, I use the peer characteristics as control variables

Table 4. Changes in peer and course composition.

Panel A. Peer characteristics.	Estimate	Std. Err.	P-value
Age	0.235	0.176	0.184
Share women	-0.009	0.005	0.069
Share with complete HS degree	-0.001	0.008	0.897
Parents' years of schooling	-0.037	0.030	0.217
Income rank by birth cohort	0.000	0.003	0.977
Days of unemployment in past year	4.755	2.895	0.102
Observations		1,989,322	

Panel B. Course characteristics.	Estimate	Std. Err.	P-value
Daytime course	-0.006	0.009	0.506
Course duration (in weeks)	0.497	0.321	0.123
Lecture hours per week	0.101	0.094	0.282
Compulsory-level course	0.001	0.003	0.809
Academic course	-0.038	0.018	0.037
Vocational course	0.038	0.018	0.037
Observations		1,989,322	

Panel C. Overall course load.	Estimate	Std. Err.	P-value
Total number of registered courses	0.117	0.129	0.367
Total lecture hours per week	0.779	0.643	0.227
Observations		415,444	

Notes: The first column reports the difference-in-differences estimate when using various peer characteristics or course characteristics as the outcome variable. In panel A and B, each individual appears once per course. In panel C, each individual appears once per academic year. Standard errors are clustered at the municipal level.

in the main regressions and check whether the estimates are sensitive to the inclusion of these controls.

Likewise, the reform may change the composition of courses that a student enrolls in or the aggregate course load that they register for. In all the results for course outcomes, I include subject-level fixed effects to account for the fact that some courses may be more difficult than others and to help alleviate concerns related to changes in course composition. However, similar to the checks for changes in peer composition, I perform two robustness checks for changes in course composition: first, I use various course characteristics as the dependent variable in my main differences-in-differences model to study whether there are differential changes in course composition between the higher-expansion and lower-expansion region. Second, I control for course characteristics (e.g., duration) in the main difference-in-differences model.

Table 4 reports the results of the difference-in-differences analysis when using different peer characteristics or course characteristics as the outcome variable, and Figures A.6 through A.8 plot the corresponding event study plots. Overall, there are no major changes in peer characteristics, though the share of female peers decreases slightly in the higher-expansion region, and the pattern is somewhat similar to what we observe for school outcomes. Likewise, there is a shift into vocational courses in the higher-expansion region compared to the lower-expansion region. However, for the other characteristics, there are no meaningful changes, and the dynamic of the effects does not mirror the effects on student outcomes. All in all, this suggests that changes in peer and course composition are unlikely to drive my findings. To support this, I further show in panel F of Table A.5 that the estimates are largely unchanged when adding controls for peer composition and course characteristics to the main difference-in-differences model.

6. Heterogeneity analysis

A policy-relevant question is whether the effects of school resource shocks are stronger for students who come from disadvantaged backgrounds or who have had low achievement levels in the past. If these students have a harder time compensating for poor resources at school, they might be particularly susceptible to changes in school inputs. Indeed, previous research at the primary and secondary level has shown that students from disadvantaged backgrounds can be more sensitive to changes in school inputs and school quality than students from more advantaged backgrounds (see, e.g., Bloom and Unterman, 2014; Jackson et al., 2016; Krueger and Whitmore, 2001). I investigate whether this also applies to adult learners by performing two different heterogeneity analyses. First, I check whether the results differ by the education level of students' parents, i.e., whether one of their parents has some post-secondary education or not. Additionally, I check whether the results differ for high school graduates and dropouts. I show the results of these two heterogeneity analyses in Figure A.9 and Figure A.10 respectively. The pattern of effects is nearly identical for students with higher- versus lower-educated parents. Similarly, there are no significant differences between high school graduates and dropouts, though the point estimates are consistently slightly larger for high school dropouts.

Research on the returns to adult education suggests that women benefit from attending adult education, while men do not (see, e.g., Blundell et al., 2020). An interesting question is whether women's academic performance is more resilient to school resource shocks, or whether men and women's performance is equally affected by the level of school inputs. I investigate whether the effects differ for men and women in Figure A.11. The results suggest no significant differences.

7. Concluding remarks

One of the most enduring and contentious debates in education research is whether school inputs have an impact on student outcomes. A vast literature examines this question for primary and secondary school students, but we know less about how school resources affect the outcomes of students beyond high school age. This is an important omission, since a notable percentage of adults are enrolled in formal or informal education in most OECD countries (OECD, 2017).

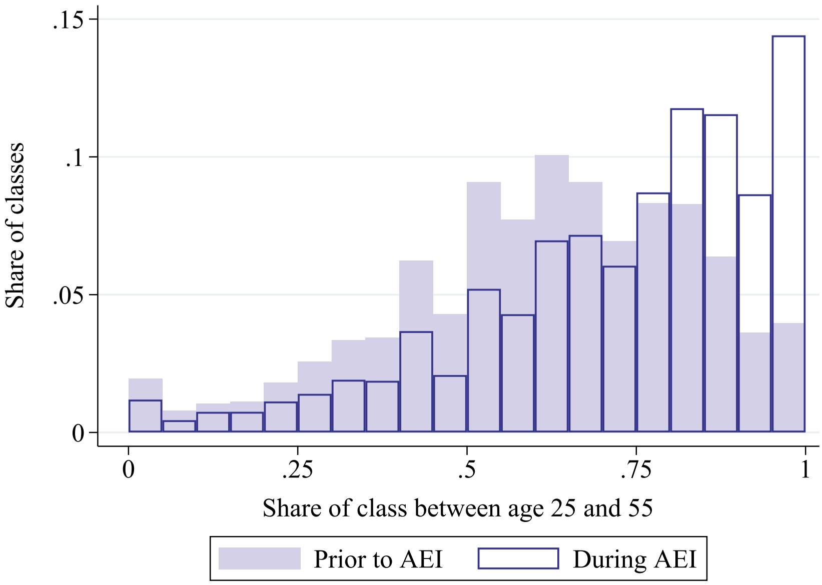
In this paper, I contribute to the literature with the first causal evidence on the relationship between school inputs and the academic outcomes of adult education students. I show that plausibly exogenous shocks to school resources such as average teacher qualifications and per-pupil expenditure coincide with increases in the probability of course dropout, though conditional upon course completion, there is no effect on course grades.

My findings suggest that policies that expand access to education without an adjustment of school inputs may have negative consequences. This is a particularly relevant finding in the context of adult education, as policymakers have begun to accept the concept of "lifelong learning" as a way to meet the changing demands of the labor market, and enrollment in adult education is on the rise. Since the shocks that I study affect multiple school inputs simultaneously, it is difficult to disentangle whether one particular input or a combination of inputs matters most. Future research could try to disentangle the mechanisms and to evaluate whether the short-term effects on course dropout have adverse consequences for students in the long run, for example on the probability to attend university or on labor market outcomes.

Appendix

Figures

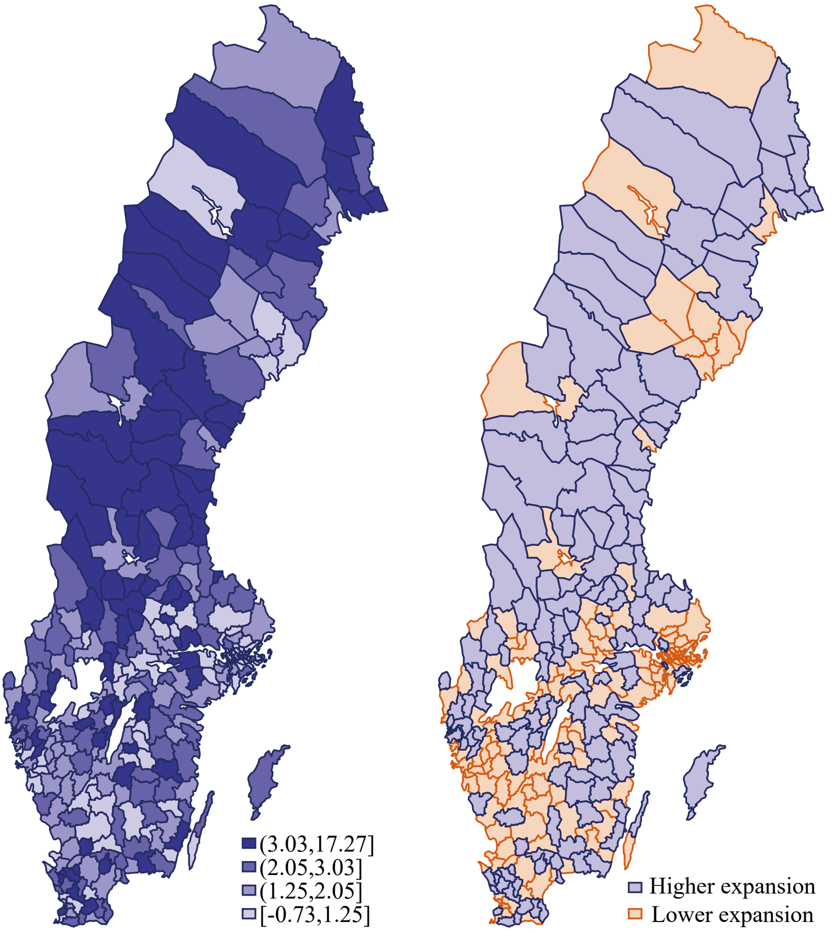
Figure A.1. Age composition of classes in municipal adult education.



Notes: This figure shows how the age composition of classes changed during the Adult Educative Initiative.

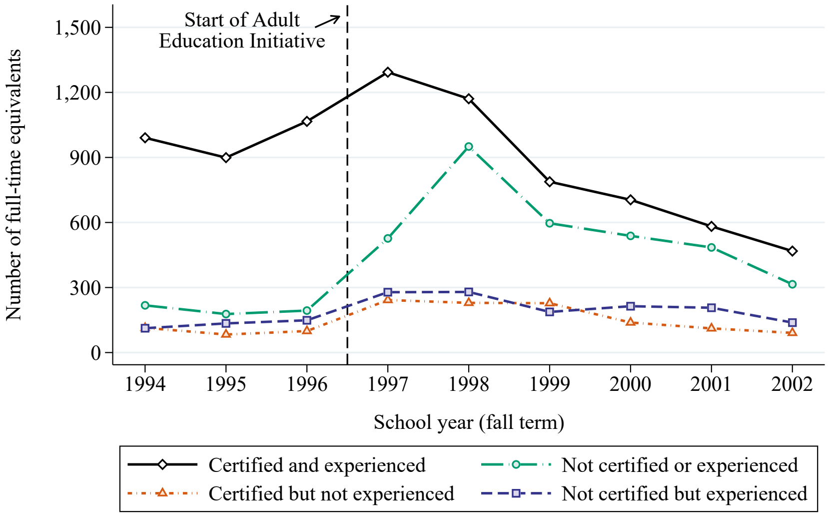
Figure A.2. Variation in the intensity of enrollment shocks across municipalities.

(a) Intensity of enrollment shock (b) Binary categorization



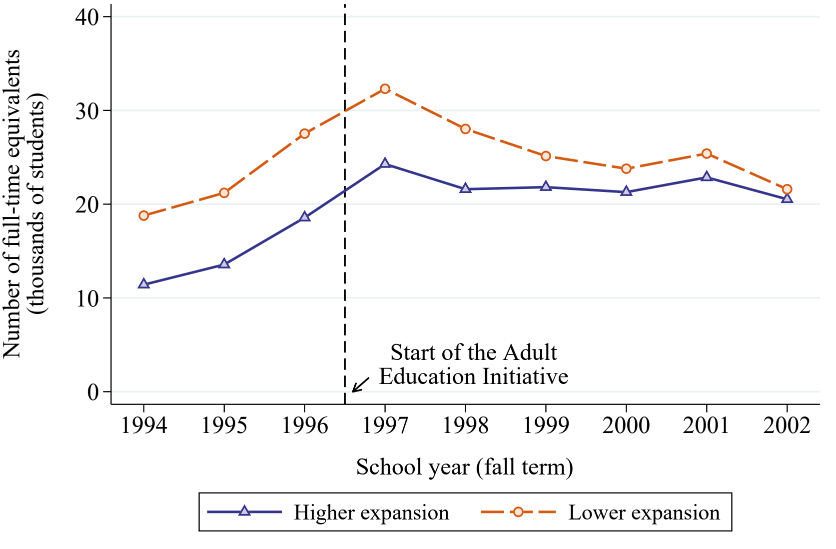
Notes: In panel (a), the darker the shade of blue, the more intense the enrollment shock. In panel (b), "higher-expansion" refers to regions that experienced above median enrollment shocks due to the AEI (shown in blue), and "lower expansion" refers to regions that experienced below median enrollment shocks (shown in orange).

Figure A.3. Inflow of adult education teachers by prior teaching experience and certification status.



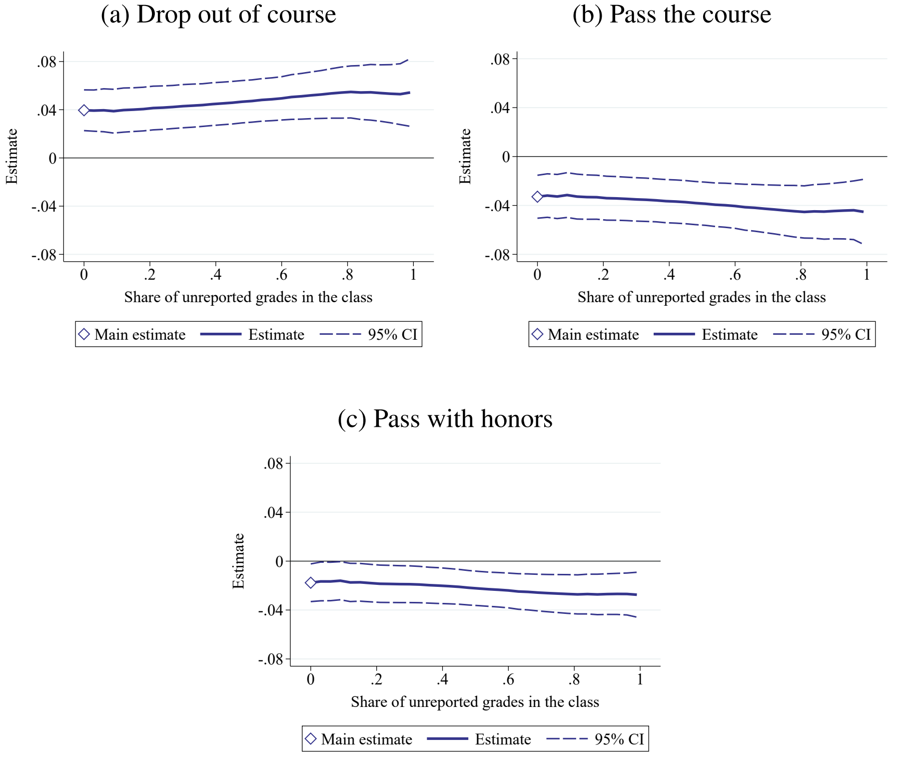
Notes: This figure shows the number of full-time equivalent teachers who taught municipal adult education during the fall term of school year y but not the previous year $y-1$. Each line categorizes teachers according to their prior teaching experience and the type of degree that they hold. Experience refers to any teaching experience since 1985, whether in adult education or another level. Certified refers to teachers who hold a college degree in pedagogy.

Figure A.4. Enrollment levels in municipal adult education (18- to 24-year olds).



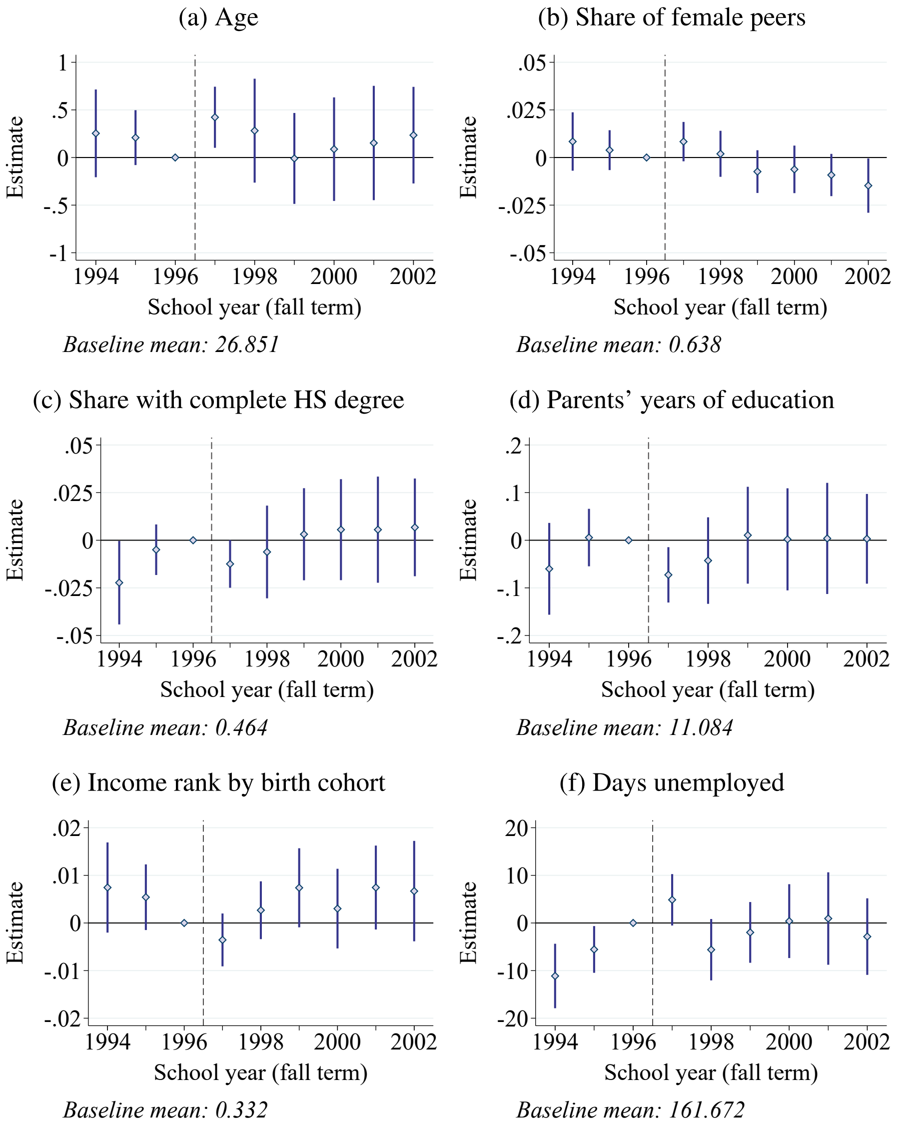
Notes: This figure shows the number of full-time equivalent students age 18 to 24 registered for municipal adult education in year *y*. The blue line plots the enrollment levels in higher-expansion regions, and the orange line plots the enrollment levels in lower-expansion regions.

Figure A.5. Sensitivity of the main estimates to inclusion of courses with different shares of unreported grades.



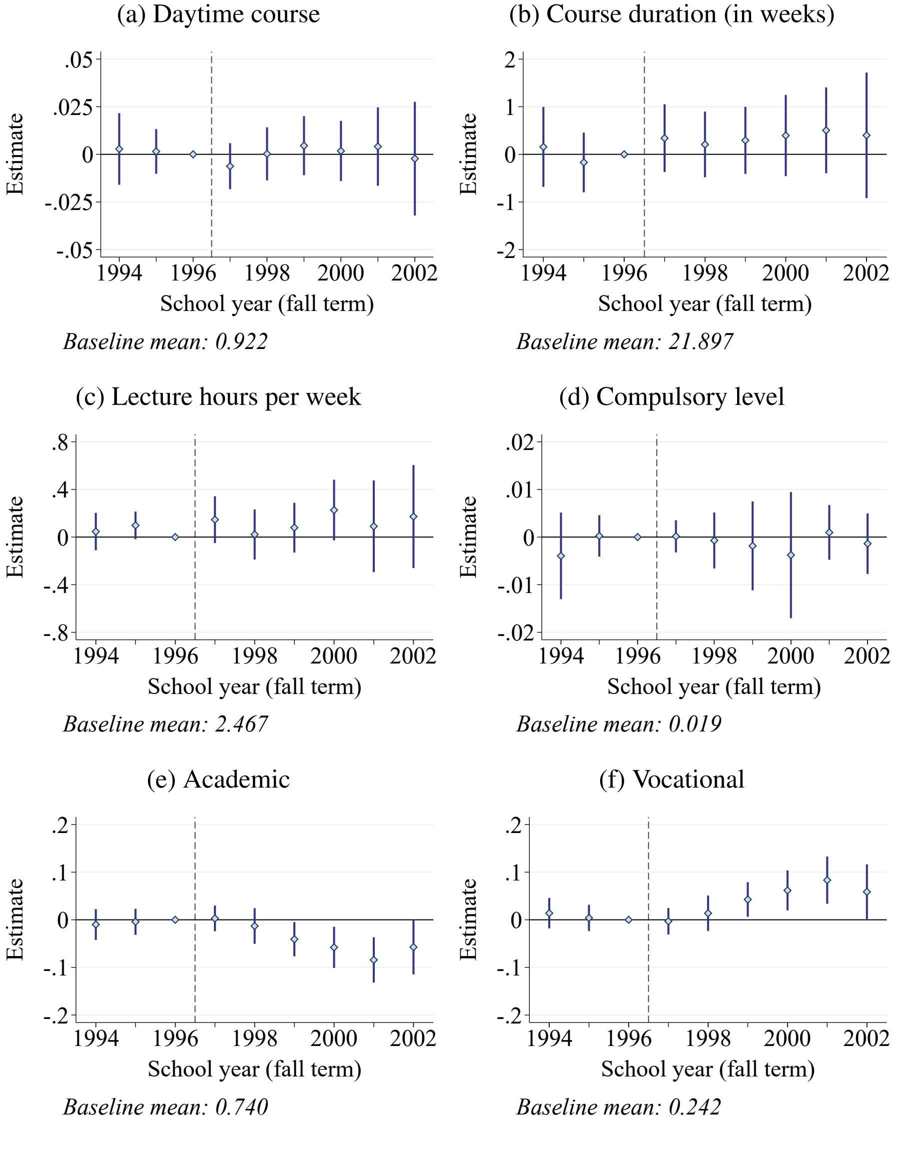
Notes: These figures show how my main estimates change when I relax the sample restriction in which I drop all courses with unreported grades. The leftmost point corresponds to the main estimates, and the further to the right along the horizontal axis, the higher the share of unreported course grades permitted for an observation to be included in the estimation. If someone has a missing grade, I treat them as a dropout in panel (a).

Figure A.6. Composition of classmates over time.



Notes: The vertical line indicates the introduction of the Adult Education Initiative. Individuals appear once for every course that they take.

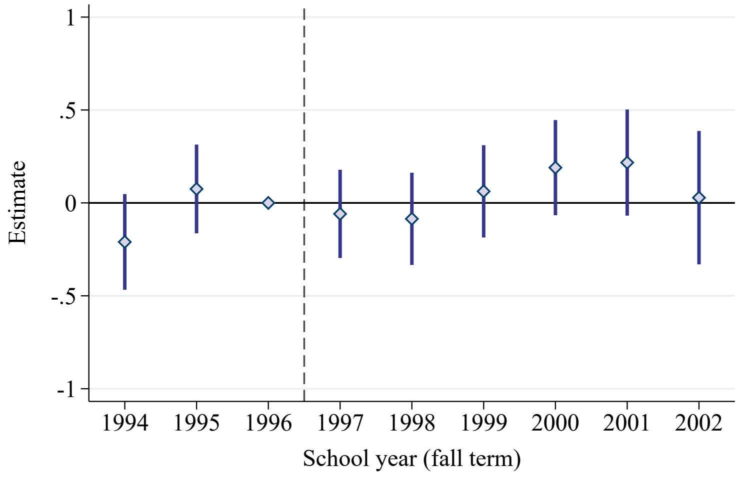
Figure A.7. Course characteristics over time.



Notes:The vertical line indicates the introduction of the Adult Education Initiative. Individuals appear once for every course that they take.

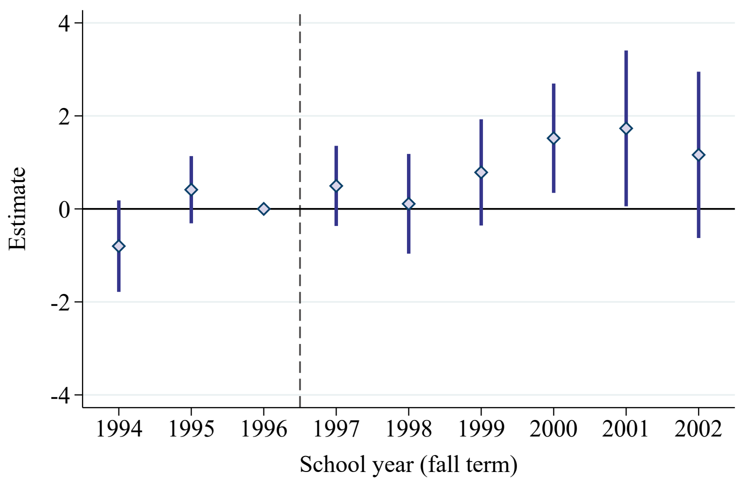
Figure A.8. Aggregate course load over time.

(a) Number of registered courses



Baseline mean: 6.855

(b) Total number of lecture hours per week

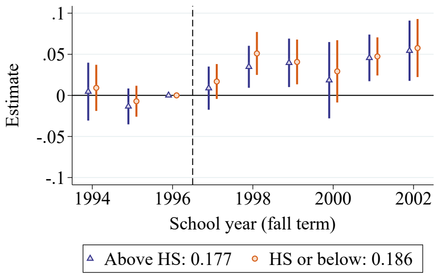


Baseline mean: 16.242

Notes: In both panels, the vertical line indicates the introduction of the Adult Education Initiative. Individuals appear once per school year.

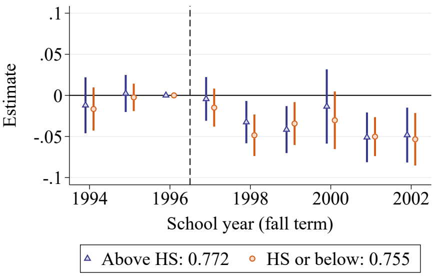
Figure A.9. Heterogeneous effects by parents' highest level of education.

(a) Drop out of course



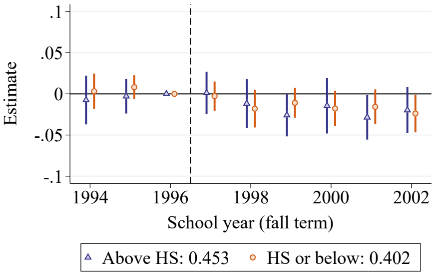
Baseline mean shown in legend.

(b) Pass the course



Baseline mean shown in legend.

(c) Pass with honors

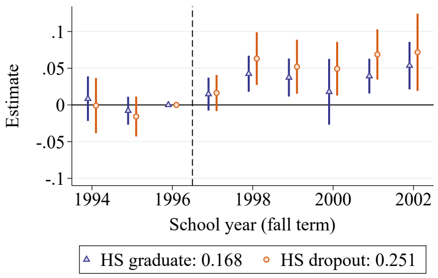


Baseline mean shown in legend.

Notes: I classify students according to the highest level of education achieved by either parent, where above high school (HS) refers to having completed any post-secondary education. The baseline share corresponds to the pre-reform mean in year 1996 in the higher-expansion regions. The vertical line indicates the introduction of the Adult Education Initiative. Individuals appear once for every course that they take. Regressions include individual controls, and the following fixed effects: subject (course prefix by level of study; municipality; and school year. Standard errors are clustered at the municipal level.

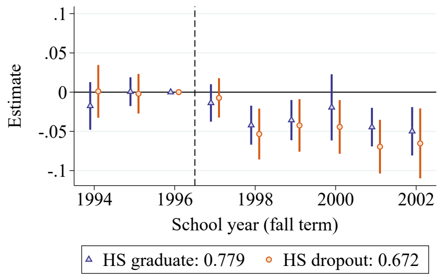
Figure A.10. Heterogeneous effects by high school graduation status.

(a) Drop out of course



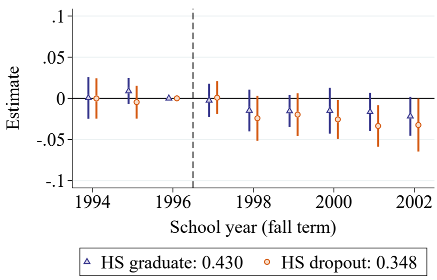
Baseline mean shown in legend.

(b) Pass the course



Baseline mean shown in legend.

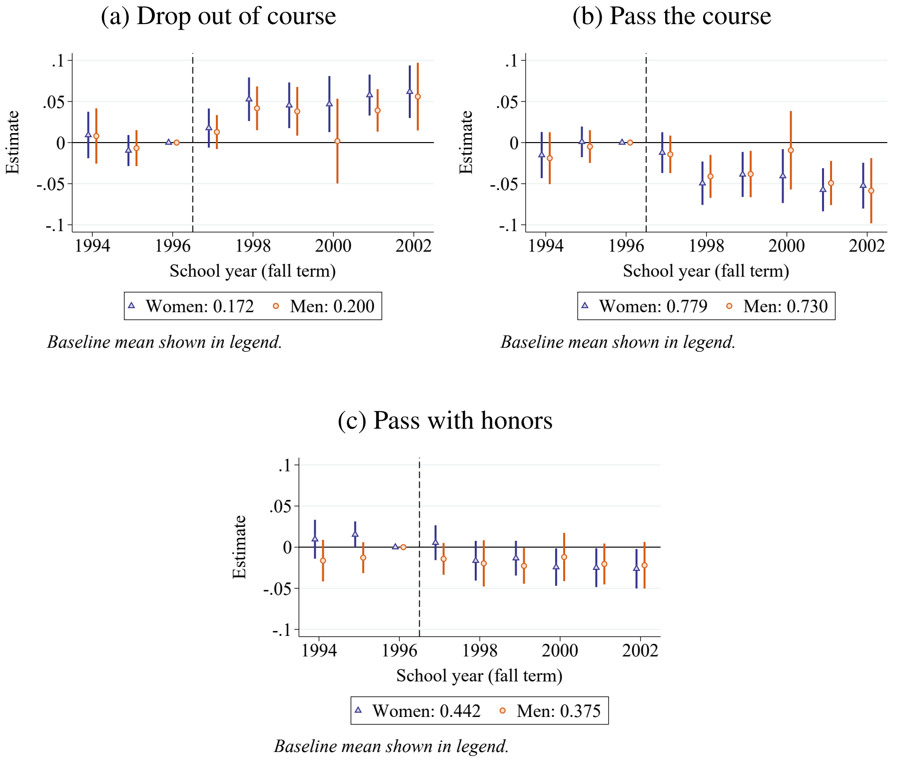
(c) Pass with honors



Baseline mean shown in legend.

Notes: A student is classified as a high school graduate if they appear in the high school graduation register; otherwise, they are a dropout. The baseline share corresponds to the pre-reform mean in year 1996 in the higher-expansion regions. The vertical line indicates the introduction of the Adult Education Initiative. Individuals appear once for every course that they take. Regressions include individual controls, and the following fixed effects: subject (course prefix by level of study; municipality; and school year. Standard errors are clustered at the municipal level.

Figure A.11. Heterogeneous effects by gender.



Notes: The baseline share corresponds to the pre-reform mean in the higher-expansion regions during the 1996 school year. The vertical line indicates the introduction of the Adult Education Initiative. Individuals appear once for every course that they take. Regressions include individual controls and the following fixed effects: municipality; school year; and subject, which is defined as the combination of a course prefix and level of study (e.g., mathematics at the upper-secondary level).

Tables

Table A.1. Descriptive statistics.

	Higher-expansion regions	Lower-expansion regions
Female	0.592	0.586
Age	21.754	21.843
	(1.794)	(1.771)
Married	0.022	0.021
Born in Sweden	0.942	0.925
Swedish-born mother	0.857	0.823
Swedish-born father	0.847	0.804
Mother's years of schooling	10.852	11.211
	(2.550)	(2.697)
Father's years of schooling	10.638	11.088
	(2.929)	(3.121)
Grade 9 GPA	-0.189	-0.106
	(0.851)	(0.828)
High school GPA	-0.334	-0.281
	(0.861)	(0.864)
Graduate of high school	0.764	0.768
Graduate of academic track	0.375	0.426
Graduate of vocational track	0.371	0.323
Graduate of other track	0.017	0.019
Observations	820,365	1,168,957

Notes: Each individual appears once per course that they are enrolled in. Higher expansion regions are those that experienced above median enrollment shocks due to the Adult Education Initiative, and lower-expansion regions are those that experienced below median enrollment shocks.

Table A.2. Descriptive statistics for outcome variables.

	Pre-reform		Post-reform	
	Higher expansion	Lower expansion	Higher expansion	Lower expansion
Panel A. School inputs.				
Class size (registrants)	40.644	43.073	34.279	38.823
	(28.440)	(27.510)	(25.843)	(27.067)
Share of certified teachers	0.855	0.864	0.742	0.796
	(0.114)	(0.115)	(0.177)	(0.181)
Share of teachers with 3+ years of experience	0.849	0.845	0.753	0.788
	(0.106)	(0.113)	(0.181)	(0.192)
Teachers' average years of schooling	15.445	15.521	15.023	15.256
	(0.528)	(0.432)	(0.729)	(0.652)
Log per-pupil expenditure on instruction	9.702	9.664	9.696	9.746
	(0.347)	(0.264)	(0.381)	(0.315)
Log per-pupil expenditure on course materials	7.309	7.197	7.258	7.295
	(0.781)	(0.766)	(0.784)	(0.791)
Log per-pupil expenditure on facilities	8.431	8.443	8.298	8.458
	(0.639)	(0.591)	(0.607)	(0.555)
Panel B. Course outcomes.				
Drop the course	0.190	0.199	0.289	0.261
Pass the course	0.747	0.732	0.652	0.665
Pass with honors	0.400	0.400	0.409	0.424

Notes: Each individual appears once per course that they are enrolled in. Higher-expansion regions are those that experienced above median enrollment shocks due to the Adult Education Initiative, and lower-expansion regions are those that experienced below median enrollment shocks.

Table A.3. Characteristics of higher-expansion and lower-expansion municipalities.

	Higher expansion	Lower expansion	Difference	P-value
A. Demographics				
Log population density (residents per square kilometer)	2.886	3.812	-0.926	0.000
Average wage income (thousands of SEK)	142.218	153.441	-11.223	0.000
Average years of schooling	10.577	10.882	-0.305	0.000
Percent female	48.320	48.799	-0.479	0.000
Percent Swedish born	90.979	89.532	1.447	0.034
Percent in workforce	69.733	73.008	-3.276	0.000
B. Industry structure				
Percent of workforce employed in:				
Agriculture, hunting, forestry and fishing	4.401	3.288	1.113	0.001
Mining and quarrying	0.410	0.411	-0.001	0.994
Manufacturing	23.495	22.858	0.637	0.567
Electricity, gas and water supply	0.999	0.743	0.255	0.026
Construction	6.591	6.135	0.456	0.010
Wholesale and retail trade	10.162	11.883	-1.721	0.000
Hotels and restaurants	2.399	2.548	-0.149	0.304
Transport, storage and communication	6.071	6.295	-0.224	0.336
Financial intermediation, real estate and business activities	7.556	9.886	-2.331	0.000
Public administration and defense	5.254	5.170	0.084	0.758
Education	6.989	6.838	0.151	0.379
Health, social work, community and personal services	25.674	23.940	1.734	0.000
Observations	143	143	286	

Notes: All characteristics are measured in 1996, the year prior to the reform. All measures other than population density are calculated for the population age 18 to 64.

Table A.4. Effects on student grades.

	Fail (1)	Pass (2)	Honors (3)
DD estimate	-0.003 (0.004)	-0.001 (0.007)	0.004 (0.008)
Pre-reform mean	0.065	0.434	0.501
Observations	1,484,769	1,484,769	1,484,769

Notes: All outcomes condition on receiving a final grade in the course. The regressions include individual controls and fixed effects for year, municipality, and subject. Standard errors are clustered at the municipal level. Pre-reform mean refers to the average in the higher-expansion regions during the 1994-1996 school years. Standard errors are shown in parentheses and clustered at the municipal level. Stars denote significance levels: *** for $p < 0.01$; ** for $p < 0.05$; * for $p < 0.10$.

Table A.5. Robustness checks for effects on student outcomes.

	Drop the course (1)	Pass the course (2)	Pass with honors (3)
Panel A. Main difference-in-differences specification.			
DD estimate	0.040	-0.033	-0.017
	(0.009)***	(0.009)***	(0.008)**
Panel B. Using continuous measure of treatment intensity.			
DD estimate	0.016	-0.012	-.009
	(0.004)***	(0.003)***	(0.009)***
Panel C. Treatment by change in FTEs from 1996/97 to 1998/99.			
DD estimate	0.029	-.023	-0.010
	(0.009)***	(0.009)***	(0.008)
Panel D. Underlying trends in municipality characteristics.			
DD estimate	0.023	-0.015	-0.008
	(0.009)***	(0.009)*	(0.008)
Panel E. Treatment by municipality of residence in 1994.			
DD estimate	0.028	-0.025	-0.012
	(0.006)***	(0.007)***	(0.006)*
Panel F. Controlling for peer and course characteristics.			
DD estimate	0.039	-0.032	-0.017
	(0.009)***	(0.009)***	(0.007)**
Pre-reform mean	0.190	0.747	0.400
Observations	1,989,322	1,989,322	1,989,322

Notes: The continuous measure used to measure treatment intensity in Panel B ranges from -0.73 to 17.27, with a median of 2.05 (see panel (a) of Figure A.2). In Panel D, I interact the time indicator with all the municipality characteristics listed in Table A.3. Standard errors are clustered at the municipal level. Pre-reform mean refers to the average in the higher-expansion regions during the 1994-1996 school years. Stars denote significance levels: *** for $p < 0.01$; ** for $p < 0.05$; * for $p < 0.10$.

Data appendix

Identifying Komvux teachers in the teacher registry

Prior to the 1999/2000 school year, it was not possible to identify the full set of Komvux teachers through the school form (SKOLFORM) variable.[17] Thus, I use the set of variables called STAD1-STAD6 in order to identify Komvux teachers for school years 1994/95 through 1998/99. Table A.6 below reports the codes that correspond to adult education teachers for these years.

Table A.6. Codes to identify Komvux teachers

School year	STAD codes
1994/95	40, 41, 42, 43, 44, 45
1995/96	34, 35, 36, 37, 38
1996/97	25, 26, 27, 28, 29
1997/98	25, 26, 27, 28, 29
1998/99	11, 12

I consider someone an adult education teacher if they have one of these codes in any of the STAD1–STAD6 variables. From the 1999/2000 school year onward, I use the level variable (NIVAKOD) that replaced the STAD1-STAD6 to identify Komvux teachers. From this year on, codes of 11 or 12 always corresponded to Komvux teachers.

Cleaning the Komvux registry

In the Komvux registry, enrollment history and course transcripts are reported at the end of each academic term. Grades are left blank for ongoing courses; for courses in which no grades are assigned (e.g., introductory courses); and for students who de-register from the course. If students do not officially de-register but fail to submit the assignments required for a final grade, teachers are supposed to record a grade of Z (*betyg underlag saknas*) on their transcript. However, in some cases, it appears that teachers have forgotten to report grades—either for the entire class or for specific individuals. On average,

[17] The issue was that the SKOLFORM variable used to be measured at the rektorsområde (principal area) level rather than the school level for all school forms except compulsory school. Because many principals organized high school and adult education in conjunction with one another, these two different school forms often existed in the same principal area. In this case, the SKOLFORM variable was always reported as high school, which meant that adult education teachers in the principal area would appear as high school teachers if the SKOLFORM variable was used for classification of teachers. This changed in the 1999/2000 update of the teacher registry: since then, different school forms in the same principal area always receive a unique code and classification.

around 15% of the grades are missing for courses where it appears that final grades should have been recorded based on course end date and student registration status. Often, grades are missing for everyone in the class, but there are also classes with only partial reporting.

In my main analysis, I take a restrictive approach and deal with missing values by dropping all classes where a student is missing a grade. However, I also show that my results are robust to different ways of dealing with the missing values. Because the Komvux registry only contains a course ID (i.e., specific to a subject) and not a class ID (i.e., specific to a group of students in the same classroom), I rely on the information contained in several other variables to identify a class. Under my definition, a class consists of anyone enrolled in the same course at the same school; furthermore, the course must have the same start and end date, the same number of lecture hours, and be held at the same time of day (i.e., daytime or nighttime). As a final restriction, I drop all introductory courses, as the vast majority (ca. 98%) do not assign any grades.

References

Andersson, C., Johansson, P., and Waldenström, N. (2011). Do you want your child to have a certified teacher? *Economics of Education Review*, 30(1):65–78.

Babcock, P., Bedard, K., and Schulte, J. (2012). No cohort left behind? *Journal of Urban Economics*, 71(3):347–354.

Bianchi, N. (2020). The Indirect Effects of Educational Expansions: Evidence from a Large Enrollment Increase in University Majors. *Journal of Labor Economics*, 38(3):767–804.

Bloom, H. S. and Unterman, R. (2014). Can Small High Schools of Choice Improve Educational Prospects for Disadvantaged Students? *Journal of Policy Analysis and Management*, 33(2).

Blundell, R. W., Salvanes, K. G., and Bennett, P. (2020). A Second Chance? Labor Market Returns to Adult Education Using School Reforms. *SSRN Electronic Journal*.

Bound, J. and Turner, S. (2007). Cohort crowding: How resources affect collegiate attainment. *Journal of Public Economics*, 91(5):877–899.

Carrell, S., Fullerton, R., and West, J. (2009). Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics*, 27(3):439–464.

Ehrenberg, R. G. and Zhang, L. (2005). Do Tenured and Tenure-Track Faculty Matter? *The Journal of Human Resources*, 40(3):647–659.

Fredriksson, P., Öckert, B., and Oosterbeek, H. (2013). Long-Term Effects of Class Size. *The Quarterly Journal of Economics*, 128(1):249–285.

Gotlands kommun (2001). Gotlands kommuns yttrande över Skolverkets tillsyn av gymnasieskola och kommunal vuxenutbildning i Gotlands kommun. [Gotland municipality's remarks on Skolverket's review of high school and municipal adult education in Gotland municipality.]. Technical report. Record number 2000:718 at Skolverket and 2000/020-60 in Gotland.

Hedges, L. V., Pigott, T. D., Polanin, J. R., Ryan, A. M., Tocci, C., and Williams, R. T. (2016). The Question of School Resources and Student Achievement: A History and Reconsideration. *Review of Research in Education*, 40:143–168.

Hoffmann, F. and Oreopoulos, P. (2009). Professor Qualities and Student Achievement. *The Review of Economics and Statistics*, 91(1):83–92.

Jackson, C. K. (2020). Does School Spending Matter? The New Literature on an Old Question. In *Confronting Inequality: How Policies and Practices Shape Children's Opportunities*, pages 165–186.

Jackson, C. K., Johnson, R. C., and Persico, C. (2016). The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. *The Quarterly Journal of Economics*, 131(1):157–218.

Jackson, C. K., Wigger, C., and Xiong, H. (2020). Do School Spending Cuts Matter? Evidence from The Great Recession. *American Economic Journal: Economic Policy*, (Forthcoming).

Jepsen, C. and Rivkin, S. (2009). Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size. *The Journal of Human Resources*, 44(1).

Krueger, A. B. and Whitmore, D. M. (2001). The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR. *The Economic Journal*, 111(468):1–28.

OECD (2017). *Education at a Glance 2017*.

Papay, J. P. and Kraft, M. A. (2015). Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. *Journal of Public Economics*, 130:105–119.

Skolverket (1998). Första året med Kunskapslyftet. [The first year of the Adult Education Initiative]. Technical report, Skolverket, Stockholm. Dnr 97:1646, delrapport nr. 3.

Stenberg, A. (2010). The impact on annual earnings of adult upper secondary education in Sweden. *International Journal of Lifelong Education*, 29(3):303–321.

III. Degree selectivity and teachers' initial job placements

Acknowledgments: This paper has benefited greatly from the input of Peter Fredriksson, Helena Holmlund, Mikael Lindahl, and Alex Solís. I am also thankful for helpful comments from David Figlio, Georg Graetz, Mark Hoekstra, Oskar Nordström Skans, Jesse Rothstein, and Björn Tyrefors, as well as seminar participants at the Institute for Policy Research at Northwestern University and the Department of Economics at Uppsala University. Additionally, I am grateful for financial support from the Jan Wallander and Tom Hedelius Foundation during my time at Northwestern University.

1. Introduction

A common aim of education policy is to ensure that all schools provide students the same learning opportunities irrespective of their background or circumstances. Equal access to qualified teachers is central to that goal given the evidence that teachers can have a substantial impact on student outcomes (Chetty et al., 2014; Rivkin et al., 2005; Rockoff, 2004). However, where teachers work, and thus who they teach, is related to their level of qualification. By virtually every measure that proxies teacher quality—experience, test scores, selectivity of undergraduate education, and certification status—less-qualified and less-skilled teachers are more likely to teach in schools with a larger share of underprivileged students (Clotfelter et al., 2005; Goldhaber et al., 2007; Lankford et al., 2002; Loeb and Reininger, 2004). There is some evidence that teacher qualification gaps overstate the true disparity in teacher effectiveness, at least according to value-added measures (Goldhaber et al., 2015; Sass et al., 2012; Steele et al., 2015). Nevertheless, it is relevant to identify the mechanisms that lead teachers with different qualifications to sort across schools in ways that disadvantage the most academically vulnerable students.[1]

This paper investigates one possible sorting mechanism that remains largely unexplored in the existing literature: that the strength of a teacher's academic credentials—in particular, the selectivity of their undergraduate degree—has a causal impact on the type of school where they find their first job. Using Chilean register data covering all college graduates and teachers from 2007 to 2020, I first motivate the analysis by documenting that teachers trained at more selective universities are unevenly distributed across schools at the start of their career. Then, I exploit features of a centralized admission process shared by a group of relatively high-prestige universities to estimate the effect of degree selectivity on teachers' initial job placements. I consider several school attributes on which teachers typically sort: public or private ownership (Baker and Dickerson, 2006; Burian-Fitzgerald and Harris, 2004); socioeconomic composition of the student body (Bonesrønning et al., 2005; Jackson, 2009; Scafidi et al., 2007); proximity to the region where they grew up (Boyd et al., 2005); and location in an urban or rural setting (Lankford et al., 2002; Monk, 2007; Thiemann, 2019).

Why might degree selectivity affect teachers' initial match to schools? Theories of statistical discrimination suggest that in order to distinguish between

[1] The most prominent research, which I have cited here, is based on data from single states in the U.S. However, there are also studies documenting that teacher qualification gaps exist in many other countries as well; see, e.g., Luschei and Jeong (2018) for an overview of international evidence.

job candidates at the start of their teaching career, employers may try to infer how well they will perform in the classroom based on a variable that is easily observable and correlated with teacher effectiveness (Altonji and Pierret, 2001). Other than years of experience, which provides no useful information in the case of novice teachers, competitiveness of undergraduate institution is one of the only observable teacher qualifications that has been linked to gains in student achievement (Clotfelter et al., 2010; Ehrenberg and Brewer, 1994; Rice, 2003). Lacking a better predictor of teacher quality, it is plausible that schools favor graduates of more selective programs when deciding which teaching candidate to hire. If so, and teaching candidates with more selective degrees prefer to work in schools serving more privileged students, this would produce the observed sorting patterns.

Despite the theoretical possibility that degree selectivity matters, few empirical studies test this hypothesis in the context of teacher labor markets. Three main challenges complicate the analysis. First and foremost, it is difficult to find random variation in degree selectivity. Graduates of more selective teaching programs differ from graduates of less selective teaching programs in important respects other than the competitiveness of their undergraduate training. These variables are likely related to their preferences for working in certain types of schools, as well as their likelihood to receive job offers from those schools. Since some of these variables are unobserved—or unobservable—by the researcher, standard analyses cannot disentangle their effect from the effect of graduating with a more selective degree. Second, there is limited access to data linking graduates, teachers, and workplaces. Researchers must often rely on nationally representative samples or focus on a small number of school districts for which they are able to obtain data. Third, even if such data do exist, it is rare to observe where graduates apply for jobs and which schools offer them positions. This leaves open the question of whether teacher preferences or school preferences drive any observed effects.

In this paper, I am able to overcome the first two of these empirical challenges and provide credible evidence on the causal relationship between degree selectivity and the characteristics of teachers' first jobs. Although I do not observe data on job applications or offers, I have rich administrative data on all individuals who graduate from teaching programs, as well as matched employer-employee data covering the entire population of teachers in Chile. The latter dataset includes all teachers who work in either public or private schools, an important feature because sorting across school sectors may widen the gaps in underprivileged students' access to qualified teachers. For identification, I exploit features of the admissions process at so-called traditional universities that produce as-good-as-random variation in access to more selective teaching degrees. Traditional universities are a group of 25 well-established and relatively prestigious institutions that share a centralized admission system. Using a deferred acceptance algorithm, applicants are admitted to programs strictly on the basis of their preferences, high school grades, and scores on a standardized entrance exam. The placement mechanism generates a score-based cutoff for each program: applicants with scores above the cutoff receive admission offers, while those with scores below the cutoff are placed on the waitlist. About 150 teacher education programs use this placement mechanism each year. My empirical analysis focuses on a sample of applicants who list a teaching program in their top preference and leverages the variation around the admission cutoff to their top-listed program in a regression discontinuity design. Intuitively, I compare the labor market outcomes of individuals who score just high enough to gain admission to their top program to the outcomes of those who just miss and end up in a program of lower selectivity instead. In order to deal with the fact that many people wait-listed for their top program end up enrolling in programs in another discipline, I focus on a preferred sample of applicants who list teaching as both their top and next-best alternative.

I find that applicants who just barely cross the admission threshold for their top-ranked program are significantly more likely to enroll in and graduate from teaching programs with better accreditation status and higher-achieving peers. They are also significantly more likely to enter the teaching profession after graduation: across the threshold, there is a six percentage point increase in the probability to teach within 10 years of application. However, the selection effect does not seem to arise because graduates with more selective degrees have an easier time finding jobs. Rather, it appears to be a near-mechanical effect driven by people below the threshold who end up in non-teaching programs and are therefore ineligible to teach. When restricting the analysis to the subsample of applicants who list teaching as their first and second choice, the probability to teach no longer jumps at the threshold. Thus, it is plausible that any observed differences in initial job placements among this marginal sample would be a consequence of sorting across schools rather than selection into and out of the profession.

My analysis of initial job placements suggests that degree selectivity does have an impact on the sorting of teachers to different types of schools. The sorting effects are most pronounced for school ownership. Getting into a more selective program significantly decreases individuals' likelihood of teaching in a public school at the start of their career. The preferred estimate indicates a decline of around 6.5 percentage points, a sizable effect that is fairly robust across different model specifications. In the context I study, there are two types of privately-managed schools where individuals can teach if they opt out of the public school sector: voucher schools, which are financed by government subsidies, and private schools, which rely exclusively on private funds. My results show that the observed decrease in the probability to work in a public school is almost entirely driven by a shift into voucher schools.

In addition to sorting by school sector, I find an impact on geographical sorting patterns, though the results are somewhat mixed. Consistent with evidence that teachers prefer to work in the region where they grew up (Boyd et al., 2005), I show that the majority of teachers work in the province where they lived during high school. However, my results suggest that, if anything, attending a more selective program reduces the likelihood of teaching close to home. While these effects are statistically indistinguishable from zero, there is other evidence that degree selectivity affects sorting across regions. In particular, crossing the admission threshold for their preferred program increases an individual's probability of working in an urban school by around three percentage points. The magnitude of the estimate is quite stable regardless of model specification, and the finding holds if I instead use a continuous measure of a school's urbanicity.

Given that rural schools and public schools tend to serve students with low socioeconomic status (SES), the patterns discussed thus far suggest that graduating from a more selective program may decrease individuals' likelihood to teach the most disadvantaged students. Indeed, although the effects are imprecise and only weakly significant in some specifications, I show that crossing the admission threshold for a more selective teaching program increases the probability of working in a medium SES school at the expense of low SES students. Interestingly, however, there does not seem to be much change in the likelihood of teaching at a high SES school. This may in part reflect the increased sorting into voucher schools: voucher schools tend to serve a more mixed student body, whereas public schools almost exclusively serve students with the lowest socioeconomic status and private schools almost exclusively serve students with the highest socioeconomic status.

The findings of this study contribute to a broad literature on teacher sorting. While it is well established that qualified teachers are unevenly distributed across schools, less is known about what drives this sorting pattern. Only a few other studies attempt to answer whether the strength of teachers' academic credentials play a role in the hiring process. The results have been inconsistent: in early work that was descriptive in nature, Ballou (1996) found that public schools in the U.S. did not seem to have a preference for candidates from more selective universities. More recently, Boyd et al. (2011) made use of a unique database on teachers' applications to transfer between schools in

New York City and reached the opposite conclusion. They found that public schools prefer to hire applicants with better pre-service qualifications, including applicants who had higher certification exam scores and applicants who graduated from more competitive colleges. In another recent U.S.-based study, Hinrichs (2014) conducted a field experiment in which he randomly assigned competitiveness of undergraduate institution to fictitious teacher résumés. He found some indication that public schools value college selectivity, but this was not the case for either private schools or charter schools, which are similar to voucher schools in the Chilean context.

My study most closely resembles Hinrichs (2014) in several respects. First, it focuses on the role of degree selectivity at the beginning of teachers' careers, a period when it should arguably have the largest effect on employers' decisions. Moreover, the empirical analysis is not restricted to the public school sector, nor is it descriptive in nature. There is, however, a key difference between our work. Because Hinrichs (2014) achieves causal identification via a résumé study, the résumé call-back rate is the only outcome that he can study. It is unclear how these response rates that he analyzes translate into job offers and placements. By contrast, my analysis focuses on actual employment outcomes and thus can provide more conclusive evidence on whether degree selectivity affects the teacher sorting patterns observed in practice.

The remainder of this paper has the following structure. In section 2, I give an overview of the institutional setting and briefly present evidence on the association between degree selectivity and teachers' early labor market outcomes. Section 3 explains the centralized admission process and the regression discontinuity design implicit in its features. Section 4 describes the data and the sample selections that facilitate my analysis. Section 5 assesses the validity of the empirical setup. Section 6 presents the main findings and demonstrates their robustness. Finally, in section 7, I discuss the results and conclude.

2. Institutional background

2.1 Undergraduate teaching programs

In Chile, prospective teachers must hold a professional teaching title. The majority earn their title by completing a full-length undergraduate program in pedagogy. Only one quarter of pedagogy students pursue alternate paths to qualification, including continuation programs for individuals with a degree in another field, as well as night and distance programs targeted at older individuals who want to begin a teaching career. The other 75% of pedagogy students matriculate in on-campus day programs, which are commonly scheduled to last four or five years. As in other fields of study, a large share of students who enter teaching programs—around 40%—drop out, and those who do finish take longer than expected to graduate—just over six years on average.

Graduation and curriculum requirements vary across teaching programs. A teaching practicum is not legally required as part of pre-service training, though some programs include a semester of practical experience. In terms of course content, all programs are specific to both the grade level and subject area in which students are formally trained to teach. Programs in primary education are typically generalist in nature, whereas programs in secondary education require extensive coursework in one particular discipline. Students choose their field of specialization at the time of application; if they later decide that they want to switch fields, they can reapply, or in special cases, transfer to a different program. Table A.1 lists all possible specializations in pedagogy. Although early and special education programs lead to a teaching title, I exclude them from my definition of teaching programs and teaching degrees throughout the remainder of this paper. Similarly, I exclude anyone working in early or special education from my definition of teachers. I make these exclusions primarily due to data limitations. Collectively, students enrolled in the remaining fields earn about 73% of all teaching titles.

Prior to and during the period I study, the government placed few restrictions on which individuals could earn teaching titles and which higher education institutions could confer them. Remarkably, neither the programs nor the institutions offering them had to go through an official accreditation process to guarantee their quality until a 2006 reform made accreditation mandatory at institutions where students received state funding. In this largely unregulated environment, teaching programs proliferated: keeping pace with broader increases in college enrollment, the number of individuals pursuing pedagogy degrees doubled between 2000 and 2008, at which point growth started to level out.[2] In 2010, the last admission year I use in my analysis, almost 80,000 undergraduate students were enrolled in full-time teaching programs at over 60 institutions with a wide geographical reach covering all of Chile's 13 administrative regions.[3]

[2]There was sustained growth in the number of students pursuing technical degrees in education over the entire sample period, but technical degrees do not lead to a professional title and only permit students to work as teaching assistants.

[3]During my sample period, three new administrative regions were created, bringing the total to 16. When I look at geographical sorting, I convert all administrative codes at the region-, province-, and commune-level back to their codes at the start of the sample period.

The increased availability of teaching programs came at the expense of their average quality and selectivity. Until the mid-2000s, traditional universities (*universidades tradicionales*)—a group of institutions founded before 1980—were the primary provider of teacher education. All 25 of these schools shared a unified admission system that allowed them to maintain relatively selective entrance requirements, as admission offers were—and still are—made solely on the basis of high school grades and performance on a standardized entrance exam. Their prestige was also boosted as a result of the Program for Strengthening Initial Teachers' Training (*Fortalecimiento de la Formación Inicial Docente*, or *FFID*), which ran from 1998 to 2002. During this period, two-thirds of the traditional universities received massive grants from the government with the aim of improving the quality of their pedagogy programs. As these universities remained committed to maintaining their quality and selectivity, their growth was relatively limited. By 2010, only 43% of pedagogy students were enrolled at traditional universities.

The rapid growth in enrollment in the early 2000s was instead driven by the spread of private establishments, many of which were unaccredited and had low admission standards. Some of the more selective private universities admitted students on the basis of the same college entrance exam as traditional universities. However, as seen by the score distributions in Figure A.1, the average score required to gain entry to teaching programs at private universities was around 480 points (about the 43rd percentile of all test-takers), compared to 550 points (about the 67th percentile of all test-takers) at traditional universities. Alongside private universities, an even less-selective class of private institutions known as professional institutes (*Institutos Profesionales*) operated programs in primary education. These professional institutes are more vocationally-oriented and only required a high school diploma for admission. Although they accounted for only 5% of the total enrollment in pedagogy, they opened up the teaching profession to even the least academically qualified students in Chile.

2.2 Entry to the teaching profession

Between 2007 and 2016, an average of 11,000 individuals per year obtained professional teaching titles in primary and secondary education. Around one-fifth pursued additional studies within two years of graduation: 4% started another undergraduate program, while 14% continued on to the post-graduate level, mostly in night and distance programs that could be completed while working.[4] Although I do not observe the work history of individuals employed outside the teaching sector, statistics obtained from the Ministry of Education indicate that between 75 and 95% of graduates successfully found some form of employment within two years, depending on the type of teaching title earned (Ministerio de Educación, 2018).[5] Since I have data on the entire population of teachers, I can compute the share of each graduation cohort that selects into the teaching profession. Figure A.2 shows the selection patterns over time. For most cohorts, around half of graduates have worked as a teacher within a year of earning their degree. Within two years, the share has risen to over 60%. Most of those who select into teaching do so within five years of graduation: by then, just under 80% have had a teaching contract. The selection patterns are relatively stable across graduation cohorts.

People who select into the teaching profession in Chile have the option to teach in three different school sectors: public, voucher, or private. All public and voucher schools are financed by the government and receive per-student subsidies; however, voucher schools are privately managed, whereas public schools are run by the municipal government. Private schools are not state-funded and are managed by the private sector. During the period of study, voucher schools employed the largest share of the teacher workforce at 47.1%, followed by public schools at 42.6% and private schools at 10.3%. Schools in these sectors differ across several important dimensions likely to affect where graduates end up working: the salaries they pay; average working conditions; socioeconomic composition of the student body; and procedures for recruiting staff.

In terms of salary, previous research has shown that schools offering higher wages can have an easier time attracting teachers (Figlio, 1997; Hanushek et al., 2004). At public schools, the salary schedule is strictly regulated by the Teachers' Statute of 1991 (*Estatuto Docente*). The minimum level of pay is based on the Basic National Wage (*Remuneración Básica Mínima Nacional* or *RBMN*) per contracted hour of teaching. This basic wage must then be topped up according to four factors: years of experience; additional training (e.g., post-graduate studies); extra responsibilities (e.g., management or technical support); and difficult working conditions. Voucher schools have somewhat more discretion when topping up wages, but they cannot pay teachers less than the minimum salary mandated by the Teachers' Statute (Santiago et al.,

[4]Author's own calculations using the enrollment registries.

[5]The aggregate data cover the 2009-2015 cohorts. Ranked from lowest to highest, two-year employment rates were as follows: 75% for those who specialized in physical education; 78% for art & music, religion & philosophy, and social sciences; 81% for foreign language; 89% for language; 90% for primary education and natural sciences; and 95% for mathematics.

2013, 2017). As an additional top-up, the government offers both individual and collective performance bonuses to high-performing teachers and schools in the public and voucher sectors. Private schools are ineligible for these government bonuses, and they are also subject to less regulation when setting teacher salaries, because they are only required to obey the minimum wage laws outlined in the Labor Code (*Codigo del Trabajo*). Teachers in private schools nevertheless earn the highest salaries on average. Teachers in voucher and public schools receive relatively similar pay, though the wage profile is slightly steeper over time for public school teachers.

In addition to salary, non-pecuniary job characteristics can influence where teachers work (Falch and Strøm, 2005; Hanushek et al., 2004). Some potentially important factors include a school's geographical location, working conditions (e.g., student-teacher ratio), and the socioeconomic composition of the student body. Because private schools charge tuition, they primarily attract students of high socioeconomic status who tend to be high-achievers on national assessments. Almost all private schools are located in urban settings. The average student-teacher ratio is relatively low at around 15 students per teacher. By contrast, more than half of public schools are located in rural areas, and they are typically attended by the lowest-achieving students from the least-privileged backgrounds: the vast majority come from poor households and have low-educated parents. There is somewhat more diversity in the population of students attending voucher schools, but on average, they are of middle to middle-high socioeconomic status and score around the mean achievement level on national tests. Around 85% of voucher schools are located in urban regions, and they have the highest average student-teacher ratio at 23 students per teacher.

On the other side of the hiring process, a school's recruitment procedures can affect both who submits an application and who receives job offers (Papay and Kraft, 2016). As in the case of teacher salaries, the Teachers' Statute outlines the rules for recruiting teachers and demands a highly transparent process in public schools. The recruitment procedure works as follows (Santiago et al., 2017): Announcements of vacancies must be made at least once a year and sufficiently publicized. All applications are reviewed by a commission consisting of the head of the municipal education administration department, the school director, and a teacher in the same subject area as the vacancy. The commission makes subjective rankings of the applicants based on criteria such as experience, past performance, and type of pedagogical training. In the end, the municipality's mayor appoints the position to the highest-ranked applicant who accepts an offer. By contrast, in voucher and private schools, the school director can be more involved in the recruitment process and has more discretion over who is hired.

In general, the pool of job candidates is sufficiently large for most schools to hire qualified teachers. Fewer than 10% of the teacher workforce lack the minimum qualifications (Ávalos and Valenzuela, 2016). However, recent survey data from PISA suggest that there are important differences in access to fully-qualified teachers across school sectors, regions, and subject areas. Principals at public schools, rural schools, and schools serving a high share of socioeconomically disadvantaged students are significantly more likely to report that shortages of qualified teachers hinder student learning "to some extent" or "a lot." Shortages of math and science teachers are the primary concern: on average, only 27% of principals perceived that teacher shortages had a negative impact on student learning in language, whereas this number was over 40% in both mathematics and science (OECD, 2013).

2.3 Degree selectivity and teacher sorting patterns

To motivate the main analysis, Table 1 documents the relationship between selectivity of undergraduate institution and outcomes on the teacher labor market. I define an institution's selectivity as the average University Selection Test score of everyone enrolled in a full-length teaching program at that school. I take the average score of all enrollees between 2007 and 2016, and then classify the institutions into one of three categories based on the average score. Low selectivity institutions are those where the average score is less than 500 (ca. 50th percentile of all test-takers); medium selectivity institutions are those where the average score is over 500 and less than 550; and high selectivity institutions are those where the average score is 550 or higher (ca. 67th percentile of all test-takers). As seen in panel A of the table, divisions based on test score also divide the sample along other dimensions of institutional quality. The average high school grades of students increases monotonically from the lowest to highest selectivity level, as does the share of students in accredited programs and institutions. The majority of students in medium- and high-selectivity schools graduate from traditional universities, which use the centralized mechanism I exploit for causal identification later in the paper; most of the outside options in private institutions fall in the low selectivity category. In terms of geographical distribution, the maps in Figures A.3 through A.5 show that high-selectivity institutions are primarily concentrated in the large administrative regions (i.e., Santiago, Bío-Bío, and Los Lagos), while low- and medium-selectivity institutions have a wider spread.

Table 1. Selectivity of institution and outcomes on the teacher labor market.

	Institution selectivity		
	Low	***Medium***	***High***
A. Characteristics of graduates			
Average score on university selection test	455.671	522.444	579.038
Average high school grades	5.486	5.678	5.947
Share from traditional universities	.151	.596	.833
Share from accredited institutions	.754	.993	.996
Share from accredited programs	.578	.837	.898
B. Selection into the teaching profession			
Share of graduates who teach within:			
One year of graduation	.510	.639	.641
Two years of graduation	.655	.749	.750
Three years of graduation	.721	.794	.794
C. Sorting by school sector			
Share of graduates who teach within:			
Public schools	.427	.338	.248
Voucher schools	.546	.604	.624
Private schools	.044	.083	.157
D. Sorting by student background			
Share of graduates who teach within:			
Low SES schools	.541	.437	.335
Medium SES schools	.310	.328	.315
High SES schools	.172	.272	.387
D. Sorting by student background			
Share of graduates who teach within:			
Rural schools	.162	.106	.060
Urban schools	.852	.908	.948
Home province	.768	.720	.687
Other province	.236	.287	.321
Number of graduates (2007-2016)	53,140	32,816	20,302

Notes: I measure an institution's selectivity by the average University Selection Test score of everyone who enrolls in a full-length teaching program at that institution between 2007 and 2016. Low selectivity: average score less than 500. Medium selectivity: average score between 500 and 550. High selectivity: average score 550 or higher. Institutions with fewer than 50 graduates are excluded from the calculations. About 15% of graduates have contracts at multiple schools, and these schools sometimes have different attributes; thus, categories such as rural vs. urban are not mutually exclusive and the sum of the shares may be greater than one.

For consistency with the labor market outcomes that I analyze later in the paper, I examine both selection into the teaching profession and different characteristics of the schools where recent graduates find their first teaching job. Panel B shows that relative to graduates of less selective institutions, graduates from more selective institutions tend to find their first teaching job faster, though the differences are negligible between the medium- and high-selectivity categories. Selectivity of undergraduate institution is also related to the types of schools where graduates teach at the start of their career. There are distinct sorting patterns across schools on every dimension that I examine: type of school management; socioeconomic background of the students served; and geographical location. For example, the first row of panel C shows that relative to graduates of low selectivity institutions, graduates from medium- and high-selectivity institutions are, respectively, about 9 and 18 percentage points less likely to find their first job in a public school. The differences are similar in magnitude across student background: graduates of medium- and high-selectivity institutions are, respectively, about 10 and 21 percentage points less likely to first teach in a school where the majority of students are of low socioeconomic status (panel D, row 1). Regional sorting patterns are somewhat weaker, but graduates of more selective institutions are less likely to work in rural schools and schools located in the province where they lived during high school (panel E, rows 1 and 3).

While these findings reveal a strong descriptive relationship between selectivity of undergraduate institution and early outcomes on the teacher labor market, they do not establish a causal connection. The correlation could simply reflect the fact that different types of people graduate from low- versus high-selectivity institutions. For example, graduates of high-selectivity schools tend to come from more advantaged socioeconomic backgrounds than graduates of low-selectivity schools. Socioeconomic background is likely related to graduates' preferences for working in one type of school over another, and could thus mediate the observed relationship between selectivity and labor market outcomes. In the next section, I will describe how I exploit the centralized admissions process at Chile's traditional universities to address the question of whether there is any causal relationship.

3. Centralized admissions and the empirical setup

3.1 Centralized admission process

Traditional universities operate the majority of medium- to high-selectivity college programs in Chile. Most undergraduate students at traditional universities apply and receive offers through a centrally coordinated admission system.[6] The selection process revolves around a standardized college entrance exam called the University Selection Test (*Prueba de Selección Universitaria*; PSU hereafter). The entrance exam is offered only once per year and consists of two mandatory sections—one in mathematics and another in language and communications—as well as two optional tests in social and natural sciences. In addition to high school grades, scores on the PSU test are the sole criteria for admission via the centralized system; together, these are used to compute composite admission scores for each applicant. No subjective factors, such as personal connections or performance at interviews, contribute to this admission score. However, universities can place different relative weight on high school grades and the various sections of the PSU test for each of the programs that they offer.

At the start of each year's admission process, all of the traditional universities publish a list of programs that they will offer in the upcoming academic year, as well as the number of vacancies and the different weights that will be placed on each admission criteria when calculating applicants' admission scores. Some programs list minimum score requirements for admission, but the admission cutoffs for each program are not predetermined. Rather, a deferred acceptance algorithm determines the admission cutoffs after all students have written the PSU test, received their scores, and submitted their applications.

Students write the PSU test on the same day all across the country and receive their scores only a few weeks later. Anyone who earns a score of at least 450 on the PSU test (about the 33rd percentile) is eligible to apply for programs via the centralized admission system. All applicants must submit a single application to the central authority (*Departamento de Evaluación, Medición y Registro Educacional*, or *DEMRE*). In this application, they can list up to eight choices ranked in order of preference. Each choice specifies a particular program at a particular university; i.e., it is a simultaneous choice of which school to attend and which subject to major in.

The placement algorithm then determines the admission cutoffs in the following manner. First, it computes each person's admission score for their top program. Applicants are provisionally assigned to their top program and ranked by strict decreasing order of their admission score. Everyone with a rank that exceeds the number of vacancies for the program is dropped and placed on the waitlist. Next, applicants who were dropped in the first round are assigned to their second option. Students in each program are re-ranked in

[6] Some special programs at these universities do not process admissions through the centralized mechanism. Additionally, a small number of students are admitted through special quotas.

decreasing order of their admission score, and once again, dropped and placed on the waitlist if their rank exceeds the number of vacancies. The process repeats until all students' listed options are exhausted. In the case of ties—e.g., if there are 30 slots for a program and applicants ranked 30, 31, and 32 have the same admission score—all applicants with the tied score receive offers for the program. As a result of this matching process, students are accepted to at most one program: the most preferred program for which their admission score is high enough.[7]

After students are informed of their admission outcomes, they have an initial three-day enrollment period during which they can accept or reject their placement. If they reject their admission slot, a spot opens up off the waitlist. Any student who does not receive an offer or who is not satisfied with their offer can re-apply the following year. However, until 2010, scores from previous PSU tests could not be used in a later admission year; rather, people wishing to re-apply had to re-take the exam, and only the current year's score was counted in the computation of the admission score.

3.2 Empirical setup

The objective of this paper is to credibly estimate the causal effect of graduating from a more selective teaching program on various sorting outcomes on the teacher labor market, including sorting by school sector, student background, and geographical location. This is an empirical challenge due to a classic "omitted variables" problem: the type of people accepted to more selective teaching programs likely differ from people accepted to less selective teaching programs on unobserved dimensions related to their probability of getting contracts in different types of schools. For example, people who attend more selective teaching programs may have stronger social networks in private schools and therefore have a higher chance of working there even absent attending a more selective teaching program. Thus, it would be unclear whether any observed effect of selectivity on the probability to work in a private school was truly a consequence of the fact that the person went to a more selective teaching program or was instead due to the fact that people from more selective teaching programs already have stronger social ties at private schools.

In order to overcome this problem, I rely on features of the centralized admission system described in section 3.1 to generate plausibly random variation

[7]This means, for instance, that students wait-listed for their first choice and admitted to their second choice cannot be admitted to or wait-listed for any of the programs listed in their third through eighth choice.

in access to more selective teaching programs. Each year, around 150 teaching programs make admission placements through the centralized mechanism. Almost 90% of these programs have a waitlist, which means the placement algorithm generates a binding cutoff above which some applicants receive a first-round offer and below which others do not. Since marginal applicants who just miss one admission cutoff typically get an offer from another program with a lower admission cutoff (or have to rely on less-elite outside options), those who successfully cross the admission threshold in the first round have a higher probability to attend and graduate from a program that is better in terms of peer characteristics and any other quality dimensions correlated with it. Given that individuals can neither perfectly predict the admission cutoffs nor perfectly manipulate their admission score to end up on one side of the threshold versus another, it should be "as good as random" whether they just cross and get access to a more selective program or just miss and get access to a less selective program (Lee and Lemieux, 2010).

The basic idea of my empirical method is to exploit the quasi-random variation in degree selectivity around the admission cutoffs to applicants' top-ranked teaching programs. I restrict my analysis to a sample of individuals who list a teaching program in their top preference and compare the labor market outcomes of people who are barely admitted to their top program to the outcomes of people who barely miss and instead end up in another teaching program. With local randomization around the cutoff, any observed differences in outcomes should capture the effect of access to more selective teaching degrees rather than differences in background characteristics.

Before going into technical details, consider the stylized example shown in Table 2, which lists two fictive applications submitted to the central authority. In this example, Person A's and Person B's top preference is to attend the Pontifical Catholic University of Valparaiso (PUCV) for a teaching program in language. While Person B's admission score of 635 is just high enough to get a first-round offer from the program, Person A's admission score of 630 is too low and they instead receive a first-round offer for their second choice, the lower-ranked teaching program in language at Arturo Prat University (UNAP). Informally, my empirical strategy compares the labor market outcomes of Person B, who was just barely accepted into the more selective teaching program at PUCV, to the outcomes of Person A, who just missed the admission threshold and was instead accepted at the less selective program at UNAP.

Table 2. Sample applications and first-round admission offers.

Person A's application						
School	**-**	**Program**	**Score**	**Cutoff**	**Outcome**
1. PUVC	-	Pedagogy in language	630.00	633.85	Wait-listed
2. UNAP	-	Pedagogy in language	630.00	553.40	Accepted
3. UTA	-	Pedagogy in language	-	520.10	-

Person B's application					
School	**-**	**Program**	**Score**	**Cutoff**	**Outcome**
1. PUVC	-	Pedagogy in language	635.00	633.85	Accepted
2. UNAP	-	Pedagogy in language	-	553.40	-
3. UTA	-	Pedagogy in language	-	520.10	-

Notes: This table provides a sample application and admission outcome to illustrate my identification strategy.

Formally, I implement a regression discontinuity design and estimate the following equation using pooled data on seven cohorts of first-time applicants to teaching programs:

$$Outcome_{ipt} = \tau \cdot \mathbb{1}[S_{ipt} \geq \tilde{S}_{pt}] + f(S_{ipt} - \tilde{S}_{pt}) + \alpha_{pt} + Type_i + \varepsilon_{ipt} \quad (1)$$

where S_{ipt} is individual i's composite admission score for their top teaching program p in application year t; $\tilde{S}_{pt}$ is the first-round admission cutoff for that program; $f(\cdot)$ is a flexible control function for distance to the top admission cutoff; and ε_{ipt} is a residual error term. The dependent variable, $Outcome_{ipt}$, denotes different sorting outcomes on the teacher labor market, such as binary indicators for whether individual i finds their first teaching job in a public school, low SES school, rural school, or school in the province where they lived during high school. The regression specification also includes program-by-year fixed effects (α_{pt}) and fixed effects for preference type ($Type_i$), which categorizes applicants into one of three groups based on the type and order of programs listed in their application (see section 4.2).

In equation (1), the parameter of interest is τ. Note that the indicator $\mathbb{1}[S_{ipt} \geq \tilde{S}_{pt}]$ takes the value of one if individual i receives a first-round admission offer for their top teaching program and zero otherwise. In practice, some applicants who receive first-round offers do not enroll in or complete their top program, while other applicants who do not receive first-round offers do. Thus, estimates of τ should be interpreted as "intent-to-treat" effects: they capture the effect of getting admitted to the top teaching program—in essence, having the opportunity to earn a degree that is more selective in terms of personal preferences, peer composition, and other quality dimensions—rather than the effect of graduating with a teaching title from that program. In order to deal with non-compliance and dropout, I also present some two-stage-least-squares estimates where I use the indicator for crossing the top admission cutoff ($\mathbb{1}[S_{ipt} \geq \tilde{S}_{pt}]$) as an instrument for completion of the top teaching program ($Degree_{ipt}$). More specifically, I estimate the following set of equations:

$$Degree_{ipt} = \delta \cdot \mathbb{1}[S_{ipt} \geq \tilde{S}_{pt}] + f_1(S_{ipt} - \tilde{S}_{pt}) + \alpha_{1,pt} + Type_{1,i} + \varepsilon_{1,ipt} \quad (2)$$

$$Outcome_{ipt} = \pi \cdot Degree_{ipt} \quad + f_2(S_{ipt} - \tilde{S}_{pt}) + \alpha_{2,pt} + Type_{2,i} + \varepsilon_{2,ipt} \quad (3)$$

where $Degree_{ipt}$ is a dummy variable equal to one if individual i graduates from their top teaching program. The predicted values from equation (2) are used in place of $Degree_{ipt}$ when estimating equation (3). The parameter of interest π is simply a rescaling of the reduced-form estimate τ. It measures the effect of graduating from the top teaching program—in a broad sense, earning a more selective degree—for marginal applicants who comply with the threshold-crossing instrument, averaged across all teaching programs and application years. In section 5, I discuss the identifying assumptions required for a causal interpretation.

4. Data and sample selection

4.1 Sources of data and definition of key variables

The key data source for this paper is a set of private-use files from the Chilean Ministry of Education that includes all registrations, applications, and admission placements via the centralized mechanism for the years 2004 through 2010. In the application database, I observe each individual's full set of ranked preferences, as well as their composite admission score, the first-round admission cutoff, and the first-round admission outcome (i.e., whether the individual was admitted or wait-listed) for each preference.[8] Moreover, the registration database contains rich demographic information about the applicants and their families, such as type of high school attended, household income, parents' education level, and parents' employment status and primary occupation.

[8]While I observe each individual's list of ranked preferences for programs with centralized admissions, I do not observe their preferences for outside options (e.g., private universities or professional institutes).

Using a personal identifier, I merge each individual's application to public-use registries that contain everyone who enrolls in or graduates from a higher education institution in Chile between the years 2007 and 2019. This allows me to determine whether individuals ultimately obtain a teaching title, and for the 2007 through 2010 cohorts, whether they initially or subsequently enroll in teaching programs after applying via the centralized mechanism. With the personal identifier, I link individuals to another public-use registry containing contract information for everyone who works in the primary and secondary school system between 2004 and 2020, irrespective of whether they work in public, voucher, or private schools. Unfortunately, I do not have any data on earnings, nor do I have information on where graduates work after college unless they select into the teaching profession. However, for those who work as teachers, the contracts database has a unique workplace identifier that allows me to determine the characteristics of the schools where they are employed. These include the average socioeconomic background of the students who attend the school; the region where the school is located; whether the school is situated in an urban or rural area; and whether the school is publicly or privately managed. Thus, I am able to study how prospective teachers from more versus less selective teaching programs sort across schools with different attributes.

All of the main outcome variables in my analysis are binary indicators, equal to one if a certain condition is fulfilled and zero otherwise. For example, in the first-stage equation (2), I estimate the jump in the probability to earn a teaching title from the top-listed program. Because I use admissions data through 2010 and only observe graduation data through 2019, I truncate the completion outcomes at nine years in order to ensure comparability across all admission cohorts in my sample. In other words, only applicants who finish their degree within nine years of their first application are counted as "completers," otherwise they are coded as zero for the aforementioned outcome. In a similar manner, I truncate all labor market outcomes at 10 years, because I only observe contracts data through 2020. Thus, anyone who does not have a teaching contract within 10 years of their first application is coded as a zero for selection into the teaching profession and has a missing value for all of the sorting outcomes.

Due to the fact that pedagogy students take a long time to complete their degree and find a job, my analysis of labor market outcomes focuses on the type of schools where graduates work at the very start of their teaching career. This is also the period of time when degree selectivity likely matters most to employers, because they have no information about novice teachers' work performance (e.g., references from previous workplaces). In order to define the various sorting outcomes, I retain information on the school(s) where each individual works during their first year as a classroom teacher. About 15% of the sample has a teaching contract at more than one school and can therefore work in multiple types of schools at the same time. If, for example, someone teaches at both a rural school and an urban school during their first year as a teacher, they are coded as one for both of these categories. In the sensitivity analysis, I show that none of my main results are sensitive if I instead use the characteristics of the teacher's primary workplace (i.e., the school where they teach the most).

4.2 Selecting the sample and setting up the RD design

In order to construct the sample database, I stack all applications submitted via the centralized admission process between 2004 and 2010. I drop invalid applications[9] and keep only the first application year in which the individual applies to a teaching program. I further restrict my sample to everyone who lists a teaching program as their top preference.[10] When performing the regression discontinuity analysis, I keep one cutoff per individual—the cutoff to their top-listed teaching program—and calculate the running variable as the distance to that admission cutoff. To ensure that there is enough variation in treatment around the threshold, I restrict the analysis to programs with a capacity of at least 15 students and at least 15 people on the waitlist. Finally, I exclude cutoffs for early and special education programs, despite the fact that these programs also lead to teaching titles. There are three reasons that I make this restriction: first, there is not universal coverage of preschool teachers in the contracts database; second, there is very limited information on the socioeconomic background of students taught by the preschool and special education teachers who I do observe; and third, certain features of the labor market (e.g., available outside options and how hiring decisions are made) are likely different for teachers in these fields. After making these restrictions, around 150 programs remain each year. On average, these programs are slightly easier to

[9]There are several types of invalid applications: applications to programs that do not exist (i.e., the applicant lists an invalid program code); repeat applications (e.g., the applicant applies to the same program in both the first and second preference); and applications made by people who do not fulfill the general admission criteria for the program (e.g., the program requires applicants to take the natural sciences sub-test, but the applicant did not complete the test).

[10]In the unrestricted sample of first-time applicants to teaching programs, the mean (median) age at application was 19.66 (19). About 42% of all first-time applicants listed a teaching program as their top preference. Conditional on being admitted to or wait-listed for a teaching program, 63% had teaching as a top choice.

get into than programs in other fields, as shown by the relative distributions of admission cutoffs in Figure A.6.

The full sample consists of 55,376 individuals. Using a set of mutually exclusive dummy variables, I classify these individuals into three preference types according to whether they *i*. only apply to teaching programs; *ii*. apply to teaching programs in their top two preferences, but apply to at least one non-teaching program further down their preference ranking; or *iii*. apply to a non-teaching program in their second preference. In my main regression specifications, I use these preference types to define a more flexible control function for the running variable, allowing the slope to vary above and below the admission cutoffs by preference type. Furthermore, although I show that my results are generalizable to the full sample, my primary analysis focuses on a "preferred sample" of type *i* and *ii* individuals. All applicants in the preferred sample list teaching as a next-best program; thus, in the event that they miss their top teaching program, their counterfactual outcome is more likely to be enrollment in and completion of a less selective teaching program. In a small window around the threshold, I show that there is no significant selection into teaching for this subsample. Thus, when looking at the characteristics of their first workplaces, I can be more confident that any observed jumps are due to sorting across schools, rather than selection into or out of the teaching profession.

Table A.2 reports descriptive statistics for both the full sample of type *i* through *iii* individuals and the preferred sample of type *i* and *ii* individuals. The preferred sample is largely representative of the full sample in terms of background characteristics, but as expected based on the sample definitions, the preferred sample applies to more teaching programs on average. I show two columns for each group: one reporting statistics for all applicants in the sample, and another reporting statistics for applicants who select into the teaching profession within 10 years of first application. Applicants to teaching programs are predominantly female and of low socioeconomic status. The gender imbalance is even larger among those who select into teaching. Whereas women make up almost 60% of applicants, they make up around two-thirds of those who ultimately become teachers. There are no remarkable differences in attrition patterns between the full and preferred sample.

Table 3 documents the main labor market outcomes for the "treated" and "control" groups in both the full and preferred sample. The treated group received a first-round offer for their top program, while the control group did not. In each sample, the differences between the treated and control groups are largely consistent with the sorting patterns described in section 2.3. In-

Table 3. Mean outcomes for those above and below the threshold.

	Full sample		***Preferred sample***	
	Treated	***Control***	***Treated***	***Control***
	(1)	(2)	(3)	(4)
A. Selection into the teaching profession				
Earn teaching title from top program	.572	.063	.575	.061
Earn any teaching title	.634	.350	.647	.411
Teach within ten years of application	.582	.341	.602	.395
Time to first teaching contract	6.244	6.607	6.252	6.578
B. Sorting by school sector				
Share who first teach in:				
Public school	.340	.385	.349	.380
Voucher school	.595	.572	.602	.585
Private school	.090	.063	.074	.057
C. Sorting by student background				
Share who first teach in:				
Low SES school	.441	.484	.457	.482
Medium SES school	.312	.322	.316	.330
High SES school	.278	.224	.257	.220
D. Sorting by region				
Share who first teach in:				
Rural school	.112	.120	.119	.116
Urban school	.902	.894	.895	.899
Home province	.710	.723	.693	.710
Other province	.296	.282	.313	.295
Number of teachers	10,409	12,779	5,775	8,736
Number of applicants	17,886	37,490	9,599	22,131

Notes: The full sample includes applicants who list a teaching program in their top preference, while the preferred sample is restricted to applicants who list a teaching program as both their top and next-best preference. The treated group received a first-round offer for their top teaching program (i.e., had a composite admission score on or above the admission cutoff), while the control group did not. For each set of sorting outcomes in panels B through D, the shares do not necessarily sum to one because teachers with more than one contract can work in multiple types of schools. About 15% of first-time teachers have contracts in different schools.

dividuals admitted to their top programs are slightly less likely to work in public schools; low SES schools; and schools located in their home province. However, there is not a consistent sorting pattern across rural and urban areas. Moreover, the differences for the preferred sample tend to be somewhat smaller in magnitude than the corresponding differences for the full sample.

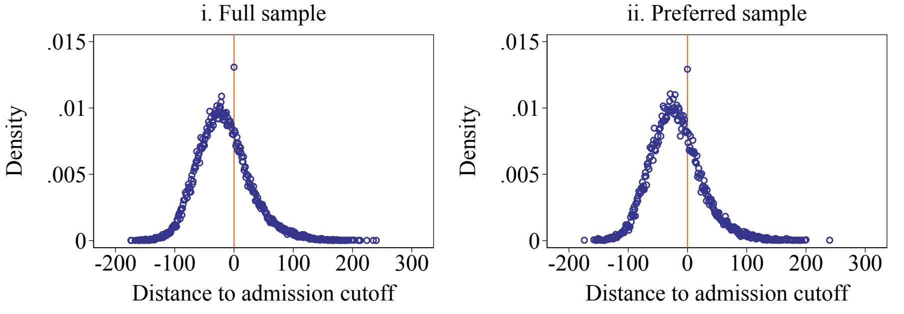
5. Validity of the RD design

5.1 Sorting across the threshold

The empirical strategy described in section 3 will only identify causal estimates of the effect of graduating from a more selective teaching program if certain conditions are fulfilled. One is that individuals should not be able to sort around admission thresholds in order to gain access to their desired program. Given the institutional setup, this is unlikely to be a realistic concern. Individuals cannot perfectly predict their top admission cutoff, as the score required for entry varies from year to year depending on the preference rankings and admission scores of all other applicants. Moreover, even if the admission cutoffs were possible to predict, it is unlikely that individuals could perfectly control the value of their admission score, because it depends largely on their performance on a centrally graded, standardized test. On the other side of the admissions process, there is not much scope for manipulation either. Programs can try to influence the type of individuals admitted to their program by adjusting the weighting scheme used to compute admission scores, but they must release these weights well in advance and cannot change them after the fact. Most importantly, a central authority is responsible for ranking applicants based solely on college entrance exam scores and high school GPA, so there is no room for manipulation based on subjective criteria or personal connections.

Despite the implausibility of sorting around the admission cutoffs, I perform several standard checks to see whether sorting seems to be a concern in practice. One is the McCrary density test, which formally tests whether there is significant bunching of applicants at the cutoff. If there is a significantly higher density of applicants located on or just above the cutoff, this would suggest that people could manipulate their scores in order to get admitted to their top program. However, it is important to note that it is not entirely unexpected to find some bunching at zero in this setup: there must always be at least one observation with zero distance to the admission cutoff for every program in every admission year, unless the individual with the lowest-admitted score is dropped under my sample restrictions. As seen in the density plots in Figure 1, this is the case here. In both the full and preferred sample, there is

Figure 1. Density plots of the running variable.



Notes: Each dot corresponds to the density of observations located within a bin of size one. The density plot on the left-hand side is for the full sample of applicants who list a teaching program as their top preference, while the density plot on the right-hand is for the preferred sample of applicants who list a teaching program as both their top and next-best choice. The vertical red line indicates the admission cutoff for the preferred teaching program.

one point right at zero with a significantly higher mass than the surrounding points. However, if the observations that lie exactly on the admission cutoff are excluded, the density of the running variable appears quite smooth, and the McCrary test finds no evidence of a statistically significant discontinuity at the threshold for almost every bandwidth between 0 and 50 (see Figure A.7).

A more direct way to test for systematic sorting is to examine whether there are any significant jumps in predetermined characteristics across the threshold. In the absence of sorting, the only thing that should jump across the threshold is access to a more selective teaching program. If there are significant jumps in predetermined background characteristics—e.g., if individuals just above the threshold have more educated parents or come from higher income families than those just below the threshold—this would suggest that the centralized admissions process does not generate as-good-as-random variation. Then, any observed jumps in outcomes on the teacher labor market could be due to jumps in other variables, not due to the jump in degree selectivity. To test for this, I estimate the main regression specification, equation (1), with different baseline characteristics as the dependent variable. The estimated jumps from these regressions are shown in Table A.3. The p-value from the joint significant test of all background characteristics is 0.265 for the full sample and 0.621 for the preferred sample, indicating that the baseline characteristics are balanced across the admission threshold. Indeed, when looking at each characteristic separately, only one of the 19 variables examined has a statistically significant jump across the threshold, which is roughly in line with the rejection rate expected simply due to random chance.

All of my main sorting outcomes condition upon having a teaching contract, so it is also important to evaluate whether there are jumps in predetermined characteristics across the threshold for the subsample of teachers. It would be concerning if, for instance, there was differential attrition of certain types of individuals just above versus just below the cutoff. This is not the case, however. Once again, the joint significance test finds no evidence of imbalance across the threshold, with a p-value of 0.473 for the full sample and 0.495 for the preferred sample. Moreover, all of the estimated coefficients are statistically indistinguishable from zero when tested separately. This reassuringly suggests that there are not important differences in some other dimension that could matter for the main outcomes, even amongst the group of individuals who select into the teaching profession after graduation.

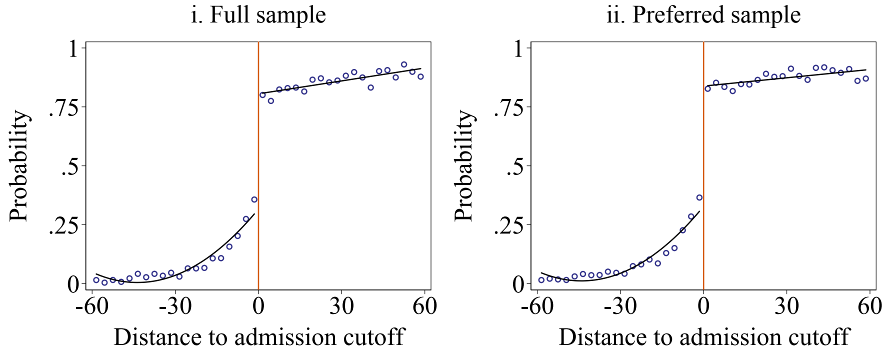
5.2 First-stage results

While background covariates should not jump at the threshold, there should be a discontinuity in the probability of entering and completing the preferred program. In other words, there should be a valid first stage. The scatterplots in Figure 2 illustrate the first-stage relationships between the running variable and the probability to enroll in and complete the preferred program. If there were perfect compliance with first-round admission offers, all dots to the left of the cutoff (indicated by the vertical line) would be equal to zero, and all dots to the right of the cutoff would be equal to one. However, because some people get accepted off the waitlist, average enrollment and completion to the left of the cutoff is greater than zero. Similarly, because some people who get into their top program decide to attend another program, and because some people who do enroll in their top program end up dropping out, average enrollment and completion to the right of the cutoff are less than one. Despite these different forms of non-compliance, there is still a very apparent jump in both enrollment and completion rates at the threshold. The regression estimates corresponding to these jumps are shown in Table 4. As seen in the first row, the likelihood of immediately enrolling in the top program increases by 50 percentage points at the cutoff. Due to the fact that some people reapply and enter their top program in a later admission year and the fact that some people drop out before earning their degree, the jumps in ever enrollment and degree completion are somewhat smaller than the jumps in immediate enrollment. However, the point estimates are still large in magnitude and statistically significant.

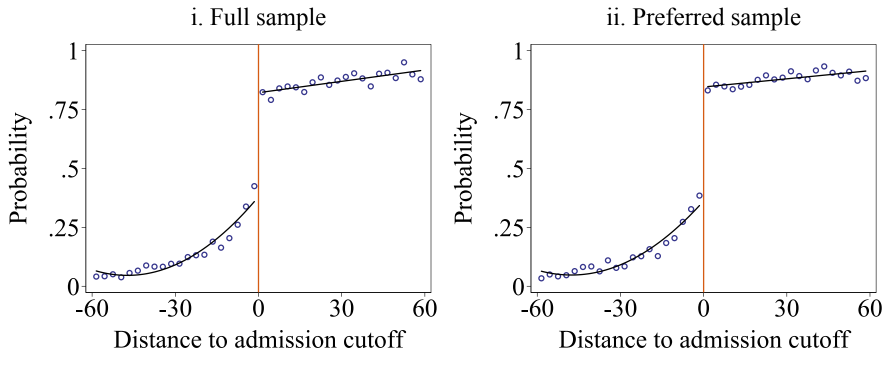
The results in Table 4 show that the threshold-crossing instrument is valid in the sense that it successfully induces a significant, discontinuous increase in

Figure 2. First-stage jumps in enrollment and completion.

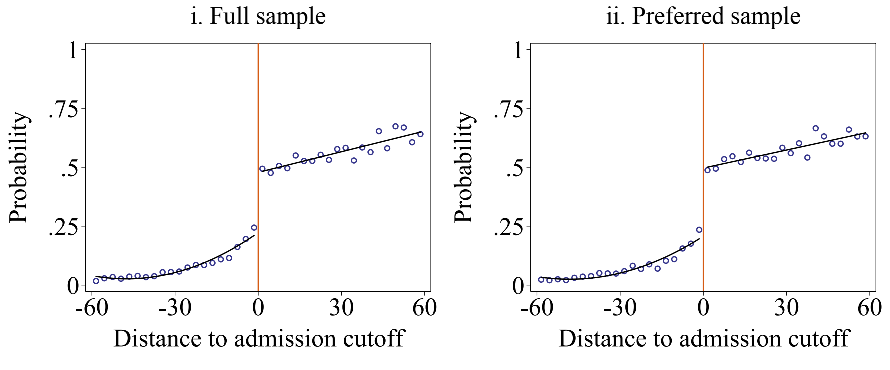
Panel A. Immediately enroll in preferred teaching program.



Panel B. Ever enroll in preferred teaching program.



Panel C. Earn a teaching title from the preferred teaching program.



Notes: The left-hand side figures are for the full sample of applicants who list a teaching program as their top choice, and those on the right-hand side are for the preferred sample of applicants who also list teaching as a next-best alternative. Each blue dot represents the average outcome within a bin of size three. The vertical orange line indicates the admission cutoff to the preferred program.

Table 4. Jumps in enrollment and completion of the preferred teaching program.

	Full sample		***Preferred sample***	
	Mean (1)	Jump (2)	Mean (3)	Jump (4)
Immediately enrolled	.146 (.353)	.500 (.013)***	.150 (.357)	.515 (.017)***
Ever enrolled	.199 (.399)	.465 (.013)***	.195 (.396)	.483 (.017)***
Earned degree	.115 (.319)	.290 (.011)***	.111 (.315)	.304 (.014)***
Observations (enrollment outcomes)	14,553		8,355	
Observations (degree outcomes)	25,709		14,718	

Notes: Columns 1 and 3 report the mean outcomes for the control group, i.e., individuals below the admission cutoff. Columns 2 and 4 report the estimated jump at the cutoff for various outcomes. The outcome "immediately enrolled" equals one if the applicant enters their preferred program the first year that they apply. "Ever enrolled" equals one if the applicant enters their preferred program within four years of first application. "Earned degree" equals one if the applicant completes their preferred program within 10 years of first application. There are more observations for completion than enrollment due to the fact that initial enrollment is unobserved for the 2004-2006 admission cohorts.

an individual's probability of enrolling in and completing their preferred program. Nevertheless, it is an open question whether people who get into their preferred program, on the margin, get access to programs that are of higher quality and that lead to better labor market outcomes for their graduates. It is relevant to shed light on this before proceeding to the main analysis. Unless getting into the preferred program in and of itself has a causal effect on labor market outcomes, it would be unexpected for threshold-crossing to lead to sorting effects in the absence of jumps in institutional and program quality.

Given that I have universal enrollment data for the years 2007 onward, I can link the last four application cohorts in my sample to the institutions and programs where they first enroll. This allows me to assess the extent to which those who cross the threshold gain access to an institution and program that is higher-ranked not only in terms of their personal preferences, but also in terms of the average achievement level of their peers and other variables that proxy for educational quality such as accreditation status. Table 5 documents which characteristics change across the threshold at the institutional level (panel A) and the program level (panel B).[11] Each row reports the estimated jumps at the cutoff when using different institutional and program characteristics as the dependent variable in the main regression, i.e., equation (1).

As expected given how the centralized admission mechanism works, multiple features of an individual's education change positively across the cutoff when they gain access to their preferred teaching program. The selectivity of a program can be thought of as encompassing all of these features at once, rather than any one in isolation. Comparing columns 2 and 4, there are relatively similar findings for both the full and preferred sample. Individuals above their preferred admission threshold are around 6 to 7 percentage points more likely to attend a traditional university. While crossing the preferred admission threshold does not increase the probability of attending an accredited institution, it raises the likelihood of attending an accredited program by about 9 to 12 percentage points, which is a relative increment of around 18% for the preferred sample and 25% for the full sample. In terms of competitiveness, there are significant jumps in the average academic achievement of the peers that applicants are exposed to. At the institution level, individuals admitted to their top program have peers who score, on average, around 14 to 15 points (around a third of a standard deviation) higher on the college entrance exam and around .07 to .09 points (also around a third of a standard deviation) higher in terms of high school grades. When instead measuring peer quality by the average grades and test scores of their classmates—i.e., people who enter the same program in the same admission cohort—the jumps are similar though a bit larger in magnitude. There are also slight changes in the educational background of peers at the institutional level. Applicants above their preferred admission threshold attend institutions with a marginally higher (lower) share of students from private (public) high schools.

Another observable aspect of an individual's education that changes across the cutoff is their field of study. As seen in panel B of Table A.5, individuals are shifted out of teaching programs in their preferred subject as their admission score falls below the cutoff for their preferred program. Figure A.8 documents these changes in more detail. It shows, for instance, that relative to individuals just above the cutoff, individuals on the waitlist are slightly more likely to specialize in primary education and slightly less likely to specialize

[11] My analysis focuses on teachers, so the ultimate interest is in how characteristics of different teaching programs change across the threshold. Thus, with the caveat that enrollment in teacher education is an endogenous outcome, I replicate Table 5 for the subsample of applicants who ever enroll in a teaching program in order to verify that there are similar jumps in institution and program characteristics conditional on selection into a teaching program. These results can be found in Table A.4 in the appendix. They are largely consistent with the findings discussed here.

Table 5. Jumps in institution, program, and peer characteristics.

	Full sample		***Preferred sample***	
	Mean (1)	Jump (2)	Mean (3)	Jump (4)
A. Institutional characteristics				
University	.916	.015	.927	.003
	(.278)	(.007)**	(.261)	(.009)
Traditional university	.651	.068	.663	.062
	(.477)	(.012)***	(.473)	(.016)***
Accredited institution	.958	.003	.963	.005
	(.200)	(.005)	(.188)	(.006)
Located in home region	.805	.017	.775	.023
	(.396)	(.013)	(.418)	(.018)
Average achievement test score	535.798	13.617	535.206	15.104
	(50.323)	(1.199)***	(46.630)	(1.594)***
Average high school grades	5.815	.065	5.808	.073
	(.230)	(.006)***	(.219)	(.007)***
Share from public high school	.367	-.011	.364	-.017
	(.129)	(.002)***	(.126)	(.003)***
Share from voucher high school	.528	-.007	.537	-.007
	(.111)	(.002)***	(.107)	(.003)**
Share from private high school	.105	.018	.099	.024
	(.119)	(.003)***	(.110)	(.004)***
B. Program characteristics				
Accredited program	.474	.118	.524	.093
	(.499)	(.013)***	(.499)	(.017)***
Duration of study (years)	4.491	.064	4.524	.074
	(.749)	(.019)***	(.684)	(.024)***
Average achievement test score	522.193	14.617	523.423	15.924
	(47.893)	(1.032)***	(46.487)	(1.397)***
Average high school grades	5.761	.087	5.767	.085
	(.234)	(.005)***	(.235)	(.007)***
Share from public high school	.396	.000	.396	-.008
	(.147)	(.003)	(.145)	(.004)**
Share from voucher high school	.537	-.006	.545	-.006
	(.127)	(.003)**	(.125)	(.004)*
Share from private high school	.066	.006	.059	.014
	(.098)	(.002)**	(.090)	(.003)***
Observations	14,225		8,172	

Notes: Columns 1 and 3 report the mean outcomes for the control group, i.e., individuals below the admission cutoff. Columns 2 and 4 report the estimates of τ from equation (1) for different institution, program, and peer characteristics. All estimates are obtained using local linear regressions with a bandwidth of 30 and a rectangular kernel. The specifications include program-by-year fixed effects, as well as a flexible control function in which the slope of the running variable can vary above and below the cutoff by preference type. Standard errors are clustered at the program-by-year level and are shown in parentheses. Significance levels are denoted with stars: *** for p-value < 0.01; ** for $p < 0.05$; * for $p < 0.10$.

in pedagogy in math and sciences. In a sense, the compositional change in the type of pedagogy degrees can be thought of as another aspect of the change in degree selectivity, because different specializations have different levels of competitiveness on average. However, the ideal comparison would arguably be made between individuals within the same specialization, given that graduates with different types of teaching titles likely apply to different teaching positions.

6. Results

In this section, I present reduced-form estimates of τ from equation (1) and second-stage estimates of π from equation (3). Unless otherwise specified, I obtain all estimates using local linear regression with a bandwidth of 30 and a rectangular kernel.[12] All regressions include program-by-year and preference-type fixed effects. Additionally, they use a flexible control function f that allows the slope of the running variable to vary above and below the threshold according to preference type.

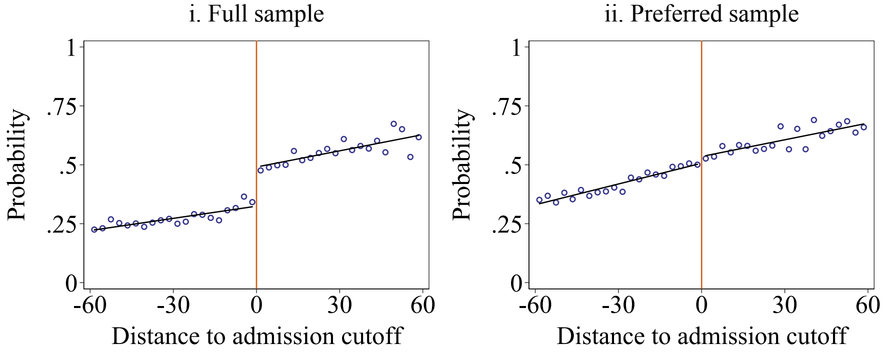
6.1 Main findings

I start by examining whether there is selection into teaching across the threshold. Figure 3 provides a graphical representation. For the full sample, there is a clear discontinuity in the probability to teach across the threshold, whereas there is no apparent jump for the preferred sample. The corresponding regression estimates are shown in row 2 of Table 6. As suggested by the scatterplot, there is indeed a significant increase in the probability to teach across the threshold for the full sample. People who cross their top admission cutoff are six percentage points more likely to have a teaching contract within 10 years of application, which is a relative increase of about 16% over the baseline mean. However, this selection effect appears to be driven by people who end up in a college major other than pedagogy. When restricting the analysis to the preferred sample of applicants who also list teaching as their next-best alternative, the estimated jump in the probability to teach drops to 1.5 percentage points, which is statistically indistinguishable from zero.

The zero-selection result for the preferred sample motivates the focus on this subgroup when turning to the sorting results, because it is less plausible

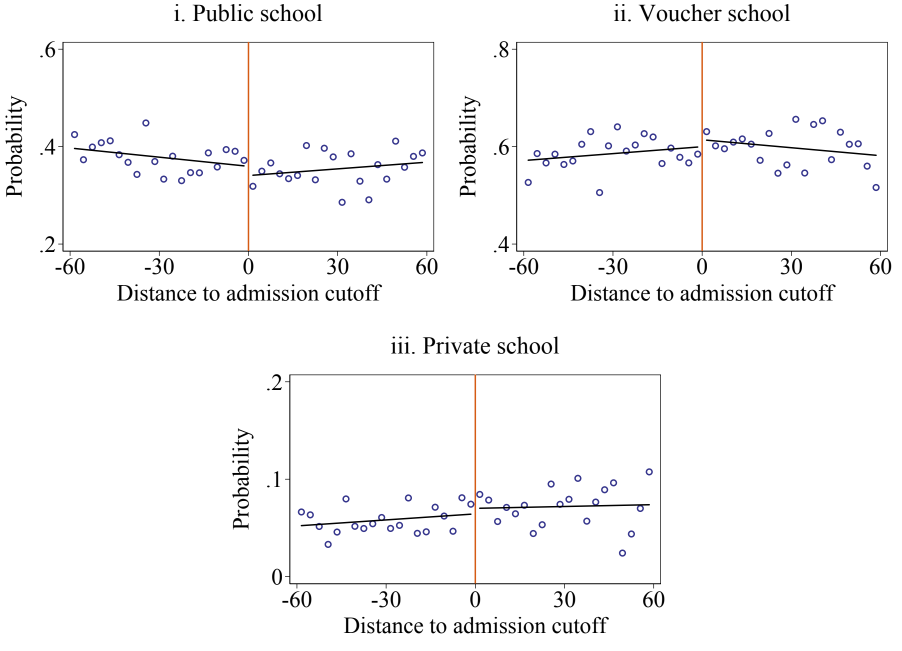
[12] In the reduced-form regressions, the optimal bandwidth according to the procedure described in Calonico et al. (2014) ranges from just under 20 to just over 30 depending on the outcome that I examine. In the sensitivity analysis, I show the main results at a wide range of different bandwidth choices.

Figure 3. Selection into teaching in the full and preferred sample.



Notes: The figure on the left-hand side shows selection into teaching for the full sample, and the right-hand side panel shows selection into teaching for the preferred sample. Each blue dot represents the average outcome within a bin of size three. The vertical line indicates the admission cutoff for the preferred teaching program.

Figure 4. Sorting by school sector in the preferred sample.



Notes: All panels depict sorting by school sector for the preferred sample. Each blue dot represents the average outcome within a bin of size three. The vertical line indicates the admission cutoff for the preferred teaching program.

Table 6. Selection and sorting in the teaching profession.

	Full sample			***Preferred sample***		
	Mean	RF	SS	Mean	RF	SS
	(1)	(2)	(3)	(4)	(5)	(6)
A. Selection into teaching						
Hold a teaching title	.399	.112	—	.481	.044	—
	(.490)	(.012)***		(.500)	(.016)***	
Work as a teacher	.392	.063	.216	.462	.015	.049
	(.488)	(.012)***	(.038)***	(.499)	(.016)	(.052)
B. Sorting by school sector						
Work in a public school	.368	-.029	-.073	.364	-.065	-.145
	(.482)	(.017)*	(.044)*	(.481)	(.022)***	(.049)***
Work in a voucher school	.583	.039	.100	.597	.057	.127
	(.493)	(.018)**	(.045)**	(.491)	(.023)**	(.051)**
Work in a private school	.068	-.004	-.011	.061	.009	.020
	(.253)	(.010)	(.024)	(.239)	(.012)	(.026)
C. Sorting by student background						
Work in a low SES school	.468	-.008	-.020	.462	-.033	-.073
	(.499)	(.018)	(.045)	(.499)	(.023)	(.051)
Work in a medium SES school	.323	.026	.065	.333	.037	.082
	(.468)	(.017)	(.044)	(.471)	(.022)*	(.049)*
Work in a high SES school	.237	-.010	-.025	.236	-.002	-.004
	(.426)	(.016)	(.039)	(.425)	(.020)	(.044)

Continued on next page.

Table 6 (continued). Selection and sorting in the teaching profession.

	Full sample			Preferred sample		
	Mean (1)	RF (2)	SS (3)	Mean (4)	RF (5)	SS (6)
D. Sorting by region						
Work in a rural school	.114	-.017	-.043	.106	-.023	-.051
	(.318)	(.011)	(.028)	(.308)	(.014)	(.032)
Work in an urban school	.897	.020	.051	.903	.028	.063
	(.304)	(.011)*	(.027)*	(.296)	(.014)**	(.030)**
Work in home province	.718	-.012	-.031	.704	-.016	-.035
	(.450)	(.017)	(.042)	(.456)	(.022)	(.048)
Work in other province	.288	.016	.041	.302	.020	.045
	(.453)	(.017)	(.042)	(.459)	(.022)	(.048)
First-stage F-stat (panel A)		941.592			602.216	
First-stage F-stat (panel B-D)		807.019			712.858	
Observations (panel A)		25,709			14,718	
Observations (panel B-D)		11,635			7,348	

Notes: Each row focuses on a different labor market outcome. Columns labeled "mean" report the mean outcome for the control group, i.e., individuals below the cutoff. Columns labeled "RF" report reduced-form estimates of the effect of getting admitted to the most-preferred teaching program (i.e., τ from equation 1). Columns labeled "SS" report second-stage estimates in which completing the most-preferred teaching program is instrumented with distance to the admission cutoff for that program (i.e., π from equation 3). All estimates are obtained using local linear regressions with a rectangular kernel and a bandwidth of 30. The specification includes program-by-year and preference-type fixed effects, as well as a flexible control function allowing the slope of the running variable to vary above and below the cutoff by preference type. Standard errors are clustered at the program-by-year level and shown in parentheses. Significance levels are denoted with stars: *** for p-value < 0.01; ** for $p < 0.05$; * for $p < 0.10$.

that any estimated effects would be driven by certain people selecting into and out of teaching. Nevertheless, for the sake of completeness and generalizability, I also present the results for the full sample. I examine several characteristics of the schools where graduates get their first teaching job: type of management (public, voucher, or private); average socioeconomic background of the students enrolled (low, medium, or high SES); and location (whether the school is situated in a rural or urban area, and whether it is located in the "home province" where the individual lived at the time of high school graduation or another province).

I begin by analyzing sorting across school sectors. Figure 4 focuses on the preferred sample and depicts how individuals' probability of finding their first teaching job in a public, voucher, or private school changes as they cross the admission threshold for their top-ranked teaching program. The pattern is a bit noisy, but as individuals cross the cutoff, we see that there is a downward jump in the likelihood of teaching in a public school, which is mirrored by an upward jump in the likelihood of teaching a voucher school. There is also a slight increase in the probability to teach in a private school, though the jump is more minute. The estimates of these discontinuities, using only the observations within 30 points of the cutoff, are shown in panel B of Table 6, under the column labelled "RF" (column 5). Consistent with the descriptive plot, there is a drop in the probability to work in a public school—a decline of around 6.5 percentage points—and a roughly equal-sized increase in the probability to work in a voucher school, indicating that the decreased sorting into public schools is primarily driven by increased sorting into voucher schools. These are large effects, especially when scaled by the first-stage jump in the probability of earning the preferred degree (column 6).

For the remaining sorting outcomes, graphical illustrations can be found in Figure A.10 in the appendix. As in the preceding example, column 5 of Table 6 reports the estimated discontinuity at the admission cutoff for each of the outcomes, while column 6 reports the second-stage estimates when I instrument completion of the preferred program with an indicator for crossing the admission cutoff.

In panel C of the table, I turn to sorting by socioeconomic background. There is suggestive evidence that graduates with more selective degrees are less likely to work in schools serving the most disadvantaged students. Crossing the admission threshold for the preferred teaching program decreases the probability of teaching in a low SES school by 3.3 percentage points, a sizable though statistically insignificant effect. This is accompanied by an increase in the probability to teach in middle SES schools, but not in high SES schools. Given the typical socioeconomic composition of schools in different sectors, this pattern is perhaps unsurprising in light of the earlier finding, i.e., the shift from public schools into voucher schools as opposed to private schools.

Finally, in panel D, I study sorting across regions. I find that crossing the admission threshold for their preferred program increases an individual's probability of teaching in an urban school by 2.8 percentage points, with a similar decrease in the probability of teaching in a rural school. I also analyze whether there is an impact on the likelihood of "returning home" to teach. If there is an effect at all, the estimates suggest that attending a more selective program may increase geographical mobility, reducing individuals' likelihood of teaching in the province where they lived during high school. However, the estimates are imprecise and insignificant.

Altogether, these results indicate that attending and graduating from a more selective teaching program has an impact on teachers' early labor market outcomes. In the following sections, I assess the robustness of these results to different model specifications, sample restrictions, and definitions of the main outcomes.

6.2 Specification checks

When implementing a regression discontinuity design, the researcher must make somewhat arbitrary choices regarding what functional form to use when controlling for the running variable; what bandwidth to select in the case of non-parametric estimation; and what weight to give each observation depending on its distance from the threshold (Lee and Lemieux, 2010). In this section, I test whether the sorting effects that I found for the preferred sample are sensitive to each of these choices.

As a first check, the graphs in Figure A.11 plot the main reduced-form estimates for the preferred sample using local linear regressions with different bandwidths ranging from five to 60. The 95% confidence intervals for the estimates are illustrated with dashed blue lines. Any time these lines enclose the horizontal zero-axis, the estimated jump is not statistically different from zero at the 5% level. At the lowest bandwidths, when few observations are used in the regressions, the estimates are quite noisy, and some are fairly sensitive in both sign and magnitude. However, in the vicinity of the optimal bandwidth (shown by the dashed orange line), the point estimates always have the same sign as the main estimate (shown by the orange dot) and are substantial in size, even when statistically insignificant.

In Table A.6, I perform additional checks on the functional form of the reduced-form specifications. The first five columns present estimates from different local linear regressions. I start with the most parsimonious regression specification in column 1, allowing the slope of the running variable to change on either side of the cutoff but including no other covariates or fixed effects. In columns 2 through 4, I sequentially add more control variables and allow for a more flexible control function. In column 2, I first add program-by-year fixed effects; then, in column 3, I replicate the main specification by adding preference-type fixed effects and interacting them with the control function so that the slope of the running variable can vary above and below the cutoff for each of the three preference types; finally, in column 4, I include an individual's age, gender, home region, and an indicator of high household income as additional controls. Reassuringly, as the regression specification becomes richer, none of the main point estimates change in meaningful ways. For one final check using non-parametric methods, I allow more weight to be placed on the observations near the cutoff by running the regressions with a triangular kernel instead of a rectangular kernel. Column 5 reports the results. The point estimates remain large in magnitude, albeit smaller than the main estimates in some cases.

Although regression discontinuity designs are local in nature, it is possible to make use of all observations in the sample rather than relying on those close to the threshold for estimation. This introduces some bias into the estimates, but can increase precision. Given that high-order polynomials have several undesirable properties (Gelman and Imbens, 2019), my final robustness check uses a quadratic form to control for the running variable. Column 6 shows that some of the estimates are sensitive to the new specification, but the main conclusions for sorting by sector and school urbanicity hold.

6.3 Different sample restrictions

As discussed in section 5, regression discontinuity designs rest on the assumption that individuals cannot manipulate their score and systematically sort on one side of the threshold. When I performed a McCrary density test to detect this kind of sorting, I found that there is in fact a significant bunching of individuals right on the admission cutoff for their top program, i.e., with zero distance to the threshold. Although this is not surprising given the institutional setup, I nevertheless perform sensitivity checks to see if the results change when I drop observations located right on the cutoff or very close to the cutoff. Colloquially, this is known as a "donut RD" design, because excluding observations on or near the cutoff creates a donut-like hole around the threshold that determines access to treatment. The results from three different donut RDs are reported in columns 2 through 4 of Table A.7. In column 2, I drop observations located exactly on the threshold, while in columns 3 and 4 I exclude all points located within one and two points of the threshold respectively. Comparing these results to the main regression results, which I replicate in column 1 for ease of comparison, there are no major qualitative differences. The magnitude of the estimates change somewhat, but again, the main conclusions hold.

Another potential sorting concern arises because individuals are permitted to re-take the college entrance exam and may postpone applying until they score high enough to get into their desired teaching program. It is possible that such individuals are different on some unobserved dimension that matters for their labor market outcomes, such as general motivation level or passion for becoming a teacher. While I have previously argued it is unlikely that people can sort perfectly on one side of the admission threshold, I address sorting concerns related to re-taking and postponing behavior by restricting the analysis to a subsample of people who submit their first application in the same year that they first take the college entrance exam. These results are reported in column 5 of Table A.7. If anything, the sorting effects become stronger.

6.4 Alternative measures of labor market outcomes

About 15% of teachers work in more than one school the first year that they teach. In the main results, I keep information on all workplaces and allow, for example, the same person to work in both a public and private school. Though this only affects a small portion of the sample, it may give a somewhat misleading picture. Thus, in column 4 of Table A.8, I show that the main results do not change when instead defining the outcome variables based on the characteristics of each teacher's primary workplace (i.e., the school where they work the most hours).

Another potential concern is that the different characteristics of schools are correlated with one another, but looking at each characteristic separately may miss important variation and sorting patterns across multiple dimensions at the same time. For example, individuals above the threshold may be less likely to teach in the most vulnerable schools—public schools in rural areas serving primarily low SES students—but looking at each characteristic separately may not adequately capture the extent of these sorting patterns. Thus, I create a composite "vulnerability index" that combines variation across multiple school characteristics. I construct the index using the first principal component of three highly correlated variables: hours worked in public school, rural school, and low SES schools (see Table A.9). I standardize the index to have a mean of zero and a standard deviation of one. Higher values indicate a more vulnerable school in terms of student population and difficulty to staff, and lower values indicate a less vulnerable school. Row 1 of Table A.10 reports the regression results from estimating equation (1) with the vulnerability index as the dependent variable. Consistent with the main results, there is a significant decrease in the composite vulnerability index at the cutoff, and the estimate is quite stable across different bandwidths (see the top panel of Figure A.12).

In a similar vein, using binary indicators for the outcome variables may miss important variation that one could exploit if a continuous measure had instead been available. Thus, for a richer measure of a school's urbanicity, I use the log population density of the commune where a school is located.[13] The second row of Table A.10 reports the results. In line with the findings that used a dichotomous measure of urbanicity, there is a significant upward jump in log population density at the cutoff, though the estimates are a bit sensitive at very low bandwidths (see the bottom panel of Figure A.12).

6.5 Heterogeneity analysis

So far, my analysis has focused on average effects. In this section, I divide the preferred sample into subgroups along two different dimensions and check whether the strength of the treatment effects differs for any particular subgroup.

The first subgroup analysis is motivated by the fact that, even amongst the traditional universities, there may be important variation in the average quality of teaching programs. Admission to the most prestigious programs in the sample may produce bigger jumps in educational quality and send a stronger signal of potential teacher quality to employers, thereby leading to more favorable labor market outcomes for graduates of these programs. In order to investigate this possibility, I create annual rankings of each teaching program in the centralized admission system according to the average college entrance exam score of all admitted applicants in a given admission year. I then categorize applicants according to whether their top-listed program is above or below median quality for their application year and run the main regression specifications separately for each subgroup. The reduced-form and second-stage estimates are reported in Table A.11. The three left-hand columns show

[13] I prefer to use the binary classification in the main results for several reasons: first, the log population density measure is only available from the 2017 census, which means it is measured years after most graduates get their first teaching job. Second, the urban vs. rural indicator from the school registry often corresponds to vulnerable schools that have the hardest time recruiting teachers, which is a relevant outcome from an equality-of-opportunity perspective.

the results for above-median programs (i.e., the most selective), while the three right-hand columns show the results for below-median programs. As hypothesized, the reduced-form estimates for sorting by school sector are notably larger for students admitted to the above-median programs, though once scaled up by the probability of degree completion, this no longer holds. All in all, the results are relatively similar across the two subsamples.

In the second subgroup analysis, I classify individuals into groups according to the type of high school they attended prior to pursuing a teaching title in order to investigate whether there are different treatment effects for public high school graduates in comparison to graduates of voucher and private high schools. Employers at voucher and private schools may have a bias that past graduates of these schools are, on average, more academically capable or better-suited to teach in environments similar to the ones they experienced as students, regardless of where they later earned their teaching title. By contrast, employers may be more uncertain about the quality of applicants who previously attended public schools and may therefore place more weight on where they earned their teaching title when assessing whether they would be a good hire. In that case, the effect of graduating from a more selective teaching program should be larger for public school graduates. Though the estimates are quite imprecise, Table A.12 suggests that there are some differences across the two subsamples. Interestingly, the increased sorting into voucher schools and medium-SES schools does not seem to be driven by public school graduates who get into their preferred program. However, in line with my hypothesis, attending and graduating from a more selective program does appear to increase the sorting of public school graduates into the most elite schools: private schools and high-SES schools.

7. Concluding remarks

A common finding in the teacher sorting literature is that teachers trained at more selective universities are unevenly distributed across schools in ways that disadvantage the most academically and socioeconomically vulnerable students. Most studies are based on data from single states in the U.S. In this paper, I use data covering the entire population of teachers in Chile and document that the same sorting pattern exists in the Chilean context. Then, I investigate whether this kind of sorting is merely a descriptive phenomenon, or whether graduating from a more selective undergraduate institution has a causal impact on the type of schools where prospective teachers find their first job. To this end, I use a regression discontinuity design implicit in the college admissions process at Chile's most selective universities and compare the early labor market outcomes of individuals on the margin of being admitted to or wait-listed for their top-ranked teaching program.

Relative to applicants just below the admission threshold, applicants who barely gain admission to their top-ranked program are more likely to enroll in and graduate from a teaching program with higher-achieving peers and better accreditation status. They are also more likely to find their first teaching job in voucher schools and schools located in more urbanized areas. Although imprecise, the estimates for sorting by socioeconomic background suggest there is also an increase in the probability of working in a medium SES school, which is mirrored by a decrease in the likelihood of working in a low SES school. Given prior research showing that teachers prefer to work close to home, I also study whether attending a more selective university affects the likelihood of working in the province where teachers went to high school. I find no indication that this is the case.

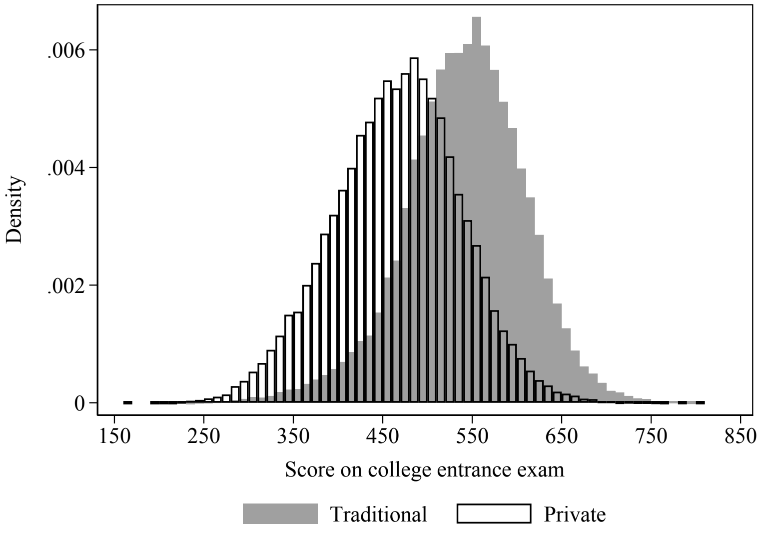
While I cannot directly observe what happens on the demand side of the hiring process, the results of this study suggest that schools value the strength of applicants' academic credentials when deciding who to hire. In particular, they appear to use competitiveness of undergraduate institution as a signal of an applicant's quality. This raises the question of the extent to which attending a more selective teaching program actually makes someone a more effective teacher, a possibility that future research could explore.

As far as the external validity of the results, Chile is arguably a context where degree selectivity is relatively likely to affect teacher sorting. The rapid expansion of teaching programs in the early 2000s led to considerable heterogeneity in the selectivity and quality of teacher training, given that many new programs were unaccredited and had low admission standards. Faced with increased uncertainty about the quality of prospective teachers—in particular graduates of newly-established programs—employers may be more dependent on degree selectivity when assessing who to hire. Moreover, Chile has an extensive voucher school system, similar to but even more pervasive than the system of charter schools in the U.S. and independent schools in Sweden. In an environment with such a high degree of school choice, school administrators likely feel extra pressure to recruit highly-qualified teachers who help them attract students.

Appendix

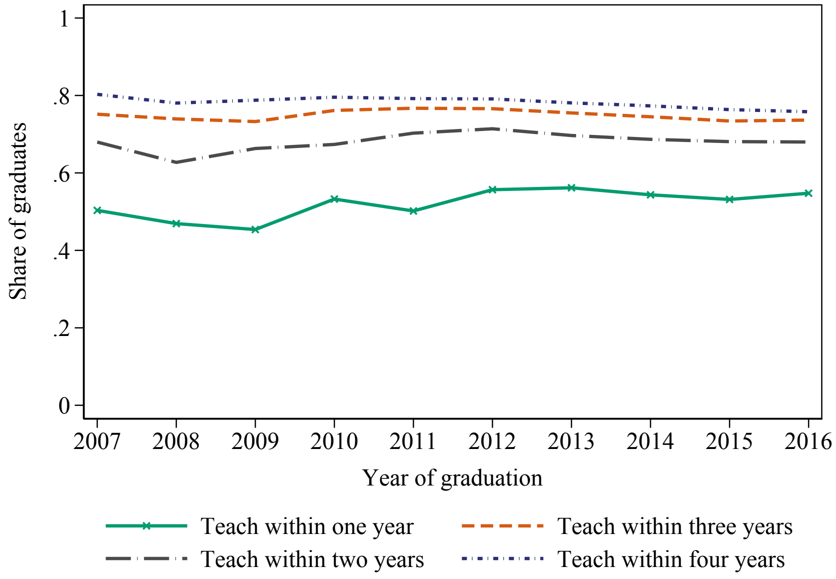
Figures

Figure A.1. Entrance exam scores for students in teaching programs at traditional universities and private institutions.



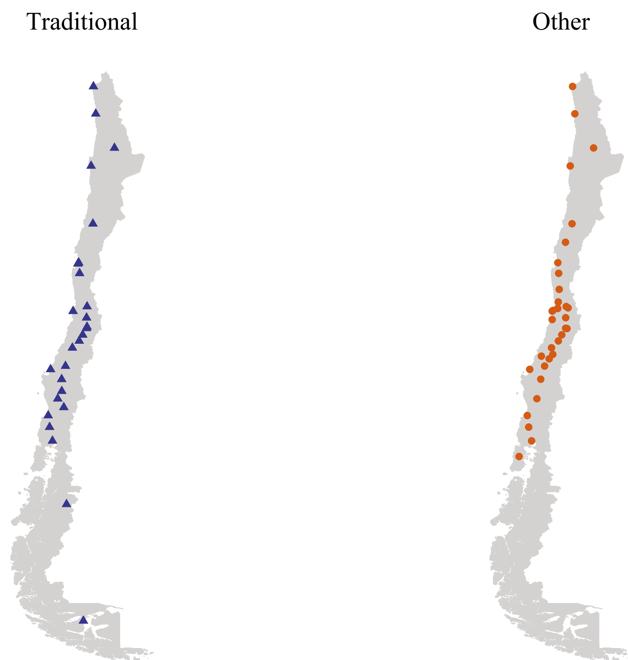
Notes: This figure illustrates the distribution of scores on the college entrance exam for students enrolled in teaching programs at traditional universities and private institutions.

Figure A.2. Selection into teaching after completion of a teaching program.



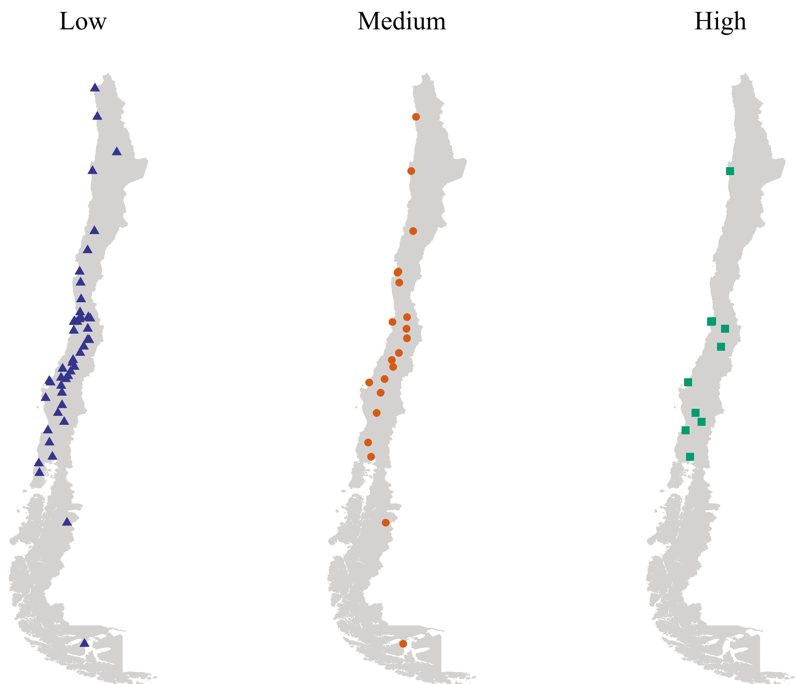
Notes: This figure plots the share of graduates between 2007 and 2016 who have worked as a classroom teacher within *t* years of earning their teaching degree. I restrict the sample to graduates of full-time teaching programs specialized in primary and secondary education.

Figure A.3. Map of Chile showing cities with students enrolled in teaching programs at traditional and other institutions.



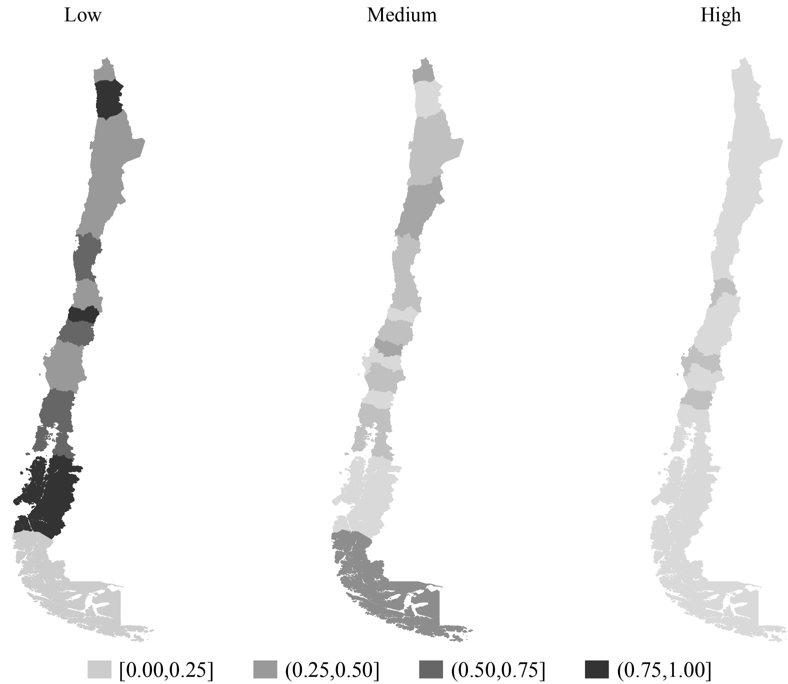
Notes: Each point marks a city where students are enrolled in full-time teaching programs specialized in primary and secondary education. The blue triangles correspond to traditional universities, and the orange dots correspond to private universities and professional institutes.

Figure A.4. Map of Chile showing cities with teaching programs at low-, medium-, and high-selectivity institutions.



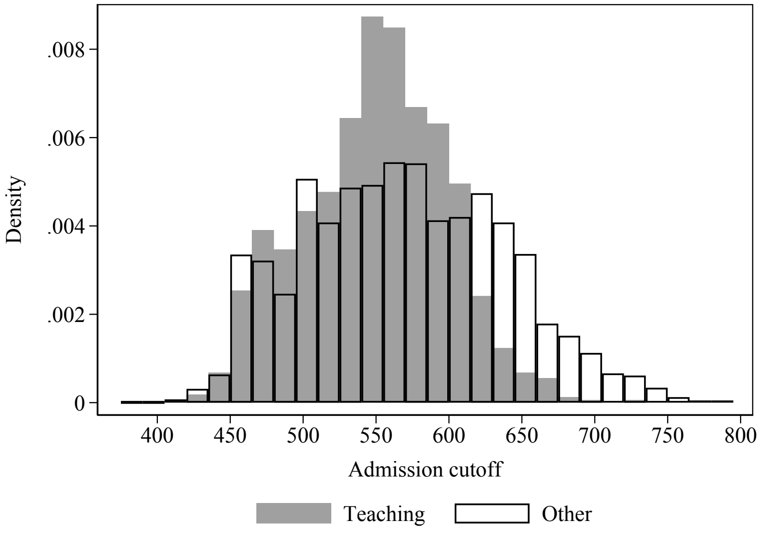
Notes: Each point marks a city where students graduate from full-time teaching programs specialized in primary and secondary education. The blue triangles correspond to low-selectivity institutions (average university selection test score less than 500); the orange circles correspond to medium-selectivity institutions (average score between 500 and 550); and the green squares correspond to high-selectivity institutions (average score 550 or higher).

Figure A.5. Share of graduates from teaching programs by selectivity of institution and administrative region.



Notes: For each of Chile's 16 administrative regions, this figure shows the share of students who graduate from teaching programs at low-, medium- and high-selectivity institutions. I define selectivity based on the average university selection test score. Low selectivity: average score less than 500. Medium selectivity: average score between 500 and 550. High selectivity: average score 550 or higher. The darker the shade, the larger the share of graduates.

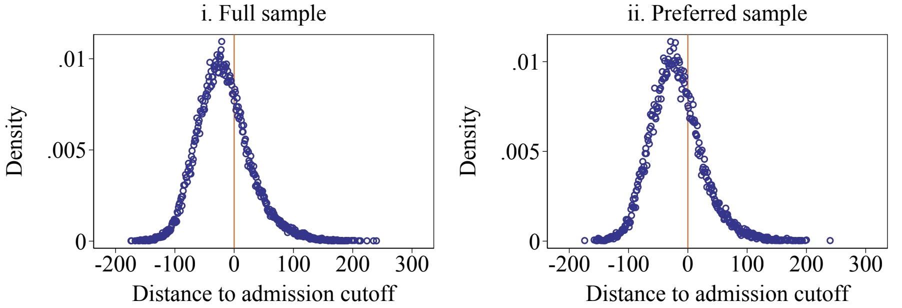
Figure A.6. Admission cutoffs for teaching programs relative to other programs.



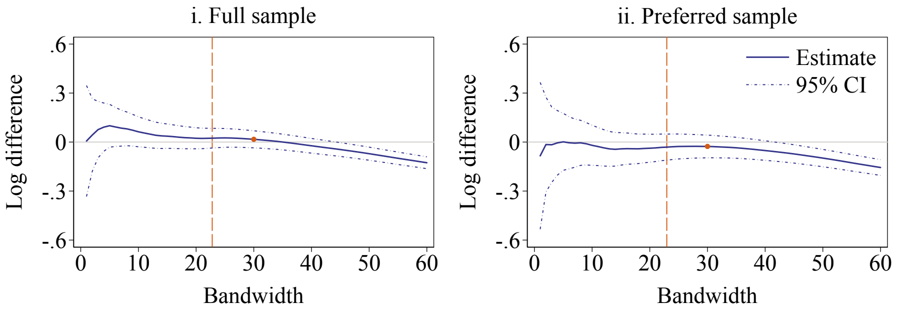
Notes: Panel A illustrates the distribution of scores on the college entrance exam for students enrolled in traditional universities and private institutions. Panel B figure illustrates the distribution of admission cutoffs for teaching programs relative to other programs for the application cohorts used in the main analysis (years 2004 through 2010). Programs that do not reach capacity are included in the figure even though the admission cutoff is not binding, which is the case for about 13.29% of teaching programs and 18.16% of other programs.

Figure A.7. Checks for discontinuities in the running variable.

Panel A. Density plot of the running variable excluding zeros.

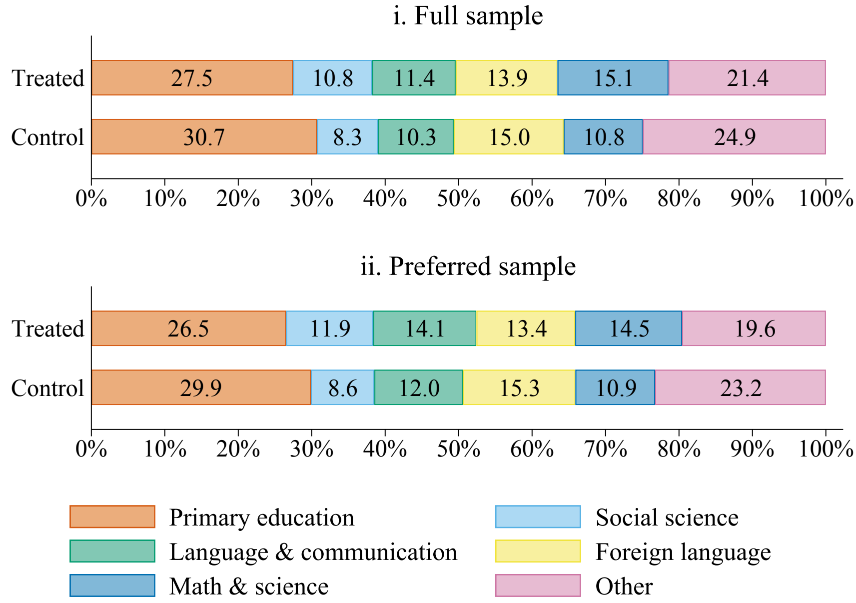


Panel B. McCrary test for jump in log difference in height at the cutoff.



Notes: Panel A shows the density of the running variable when excluding zeros (i.e., the points that lie exactly on the cutoff). Each dot corresponds to the density of observations located within a bin of size one. Panel B reports the results of the McCrary density test at different bandwidths ranging from one to 60, also when excluding zeros. The solid lines plot the estimated log difference in height at the admission cutoff, and the dashed lines show the 95% confidence intervals for the estimates. The dashed orange line marks the optimal bandwidth chosen by the selection procedure in McCrary (2008), and the orange dot at 30 marks the bandwidth used in the regression tables.

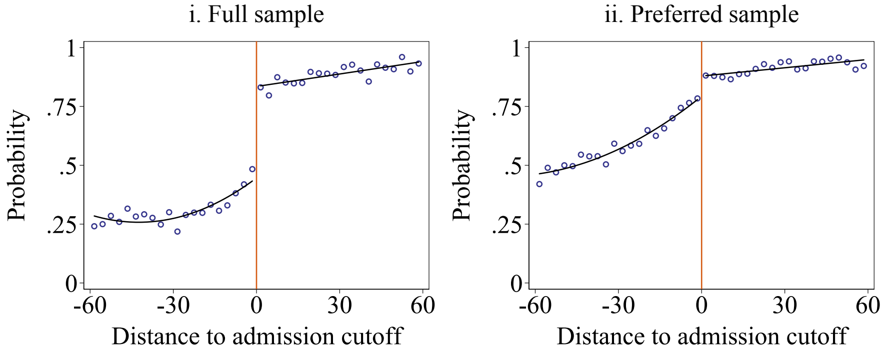
Figure A.8. Type of teaching titles earned.



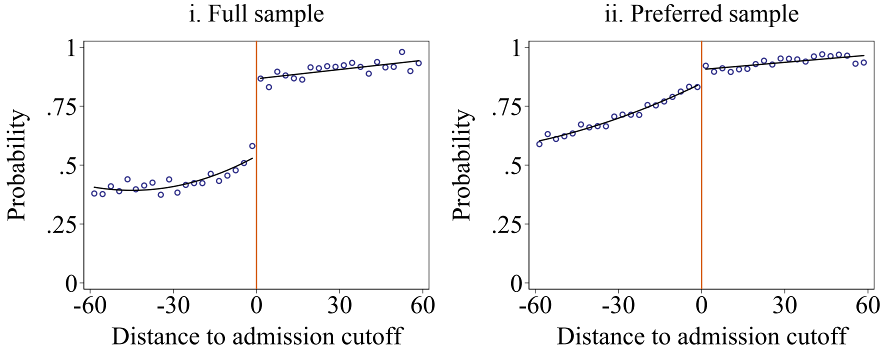
Notes: Treated refers to applicants who score above the admission cutoff and receive a first-round offer for their top-ranked teaching program. Control refers to applicants who score below the cutoff and do not receive a first-round offer for their top-ranked program. The sample is restricted to a bandwidth of 30, as in the main regressions. Other teaching titles include specializations in art and music, philosophy and religion, physical education, and technology.

Figure A.9. Jumps in enrollment and completion of any teaching program.

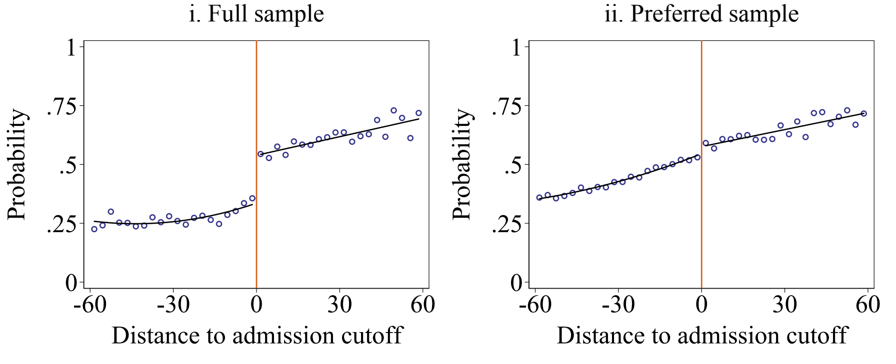
Panel A. Immediately enroll in any teaching program.



Panel B. Ever enroll in any teaching program.



Panel C. Earn a teaching title from any teaching program.

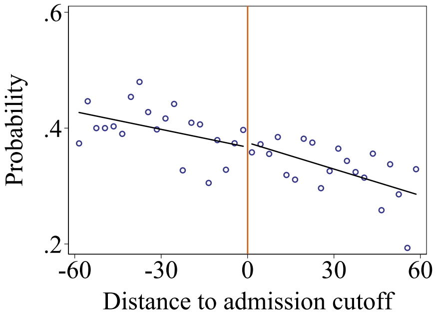


Notes: The left-hand side figures are for the full sample of applicants who list a teaching program as their top choice, and those on the right-hand side are for the preferred sample of applicants who also list teaching as a next-best alternative. Each blue dot represents the average outcome within a bin of size three. The vertical line indicates the admission cutoff to the preferred program.

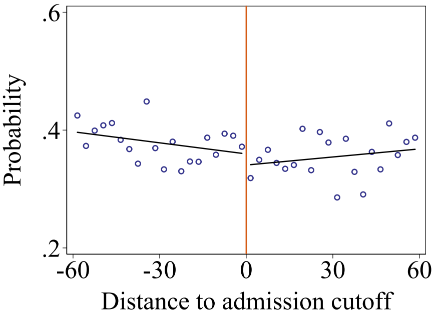
Figure A.10. Scatterplots for teacher sorting outcomes.

Panel A. First teaching job in a public school.

i. Full sample

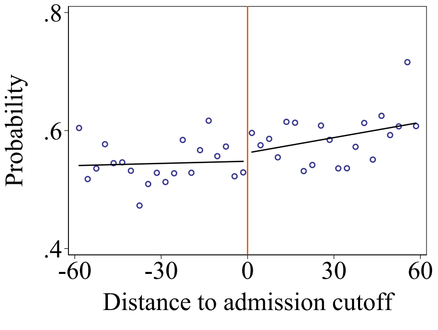


ii. Preferred sample

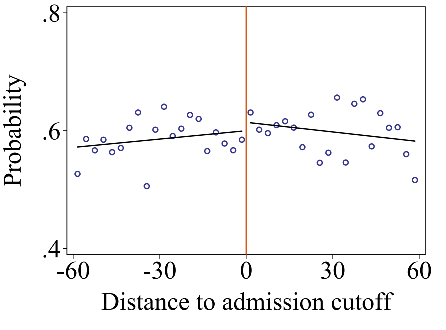


Panel B. First teaching job in a voucher school.

i. Full sample

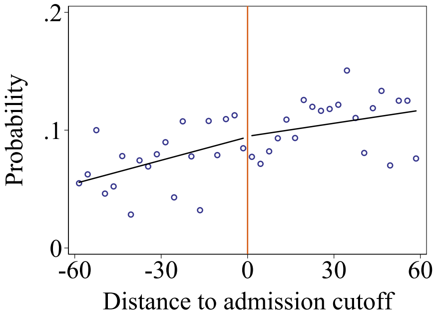


ii. Preferred sample

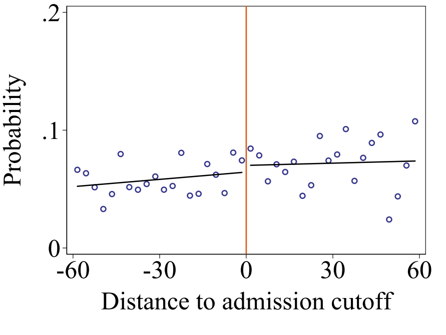


Panel C. First teaching job in a private school.

i. Full sample



ii. Preferred sample

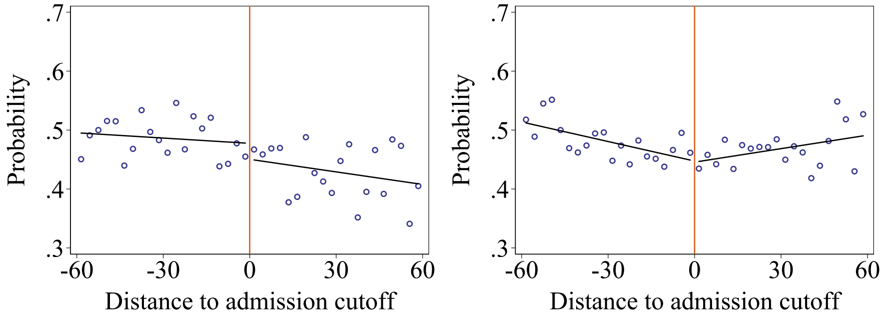


(Continued on next page.)

Figure A.10 (continued). Scatterplots for teacher sorting outcomes.

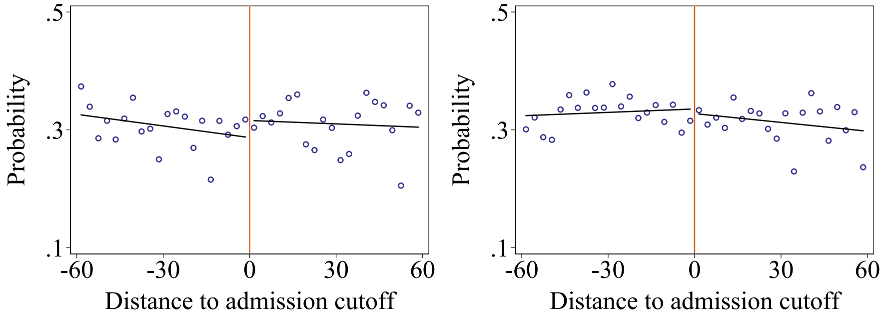
Panel D. First teaching job in a low SES school.

i. Full sample | ii. Preferred sample



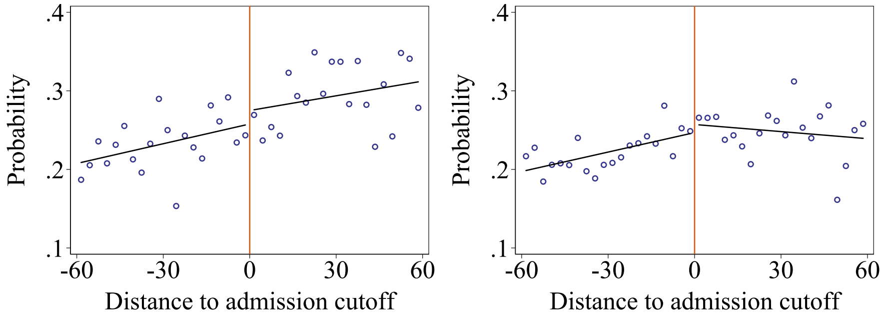
Panel E. First teaching job in a medium SES school.

i. Full sample | ii. Preferred sample



Panel F. First teaching job in a high SES school.

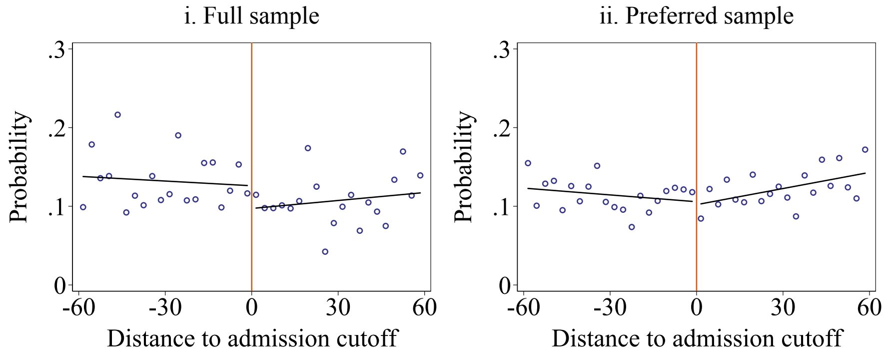
i. Full sample | ii. Preferred sample



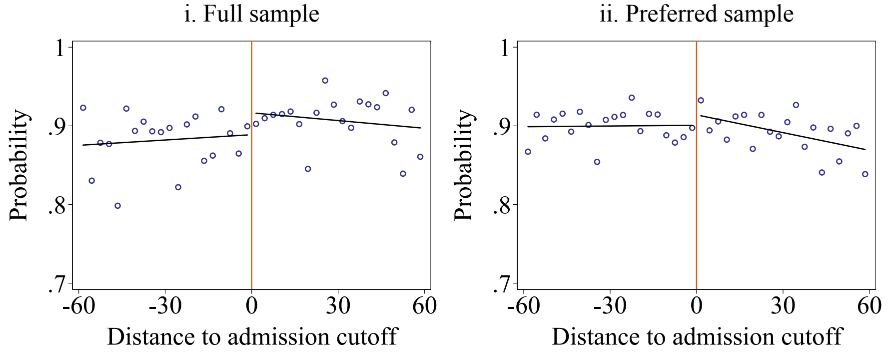
(Continued on next page.)

Figure A.10 (continued). Scatterplots for teacher sorting outcomes.

Panel F. First teaching job in a rural school.



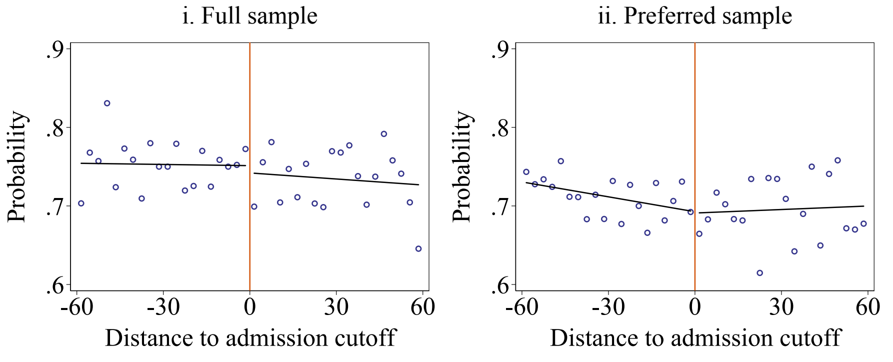
Panel G. First teaching job in an urban school.



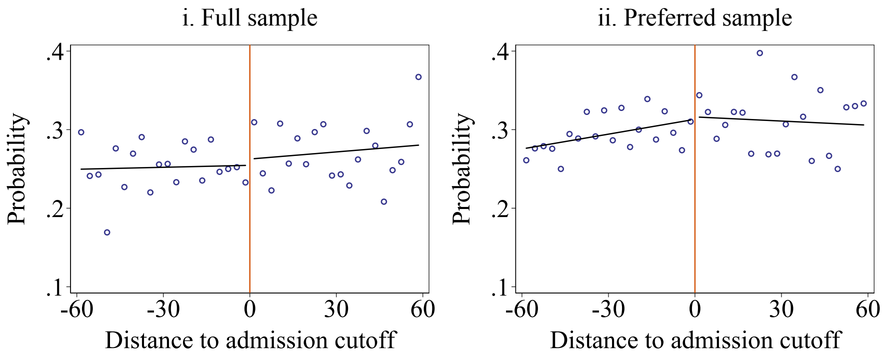
(Continued on next page.)

Figure A.10 (continued). Scatterplots for teacher sorting outcomes.

Panel I. First teaching job in a school located in home province.

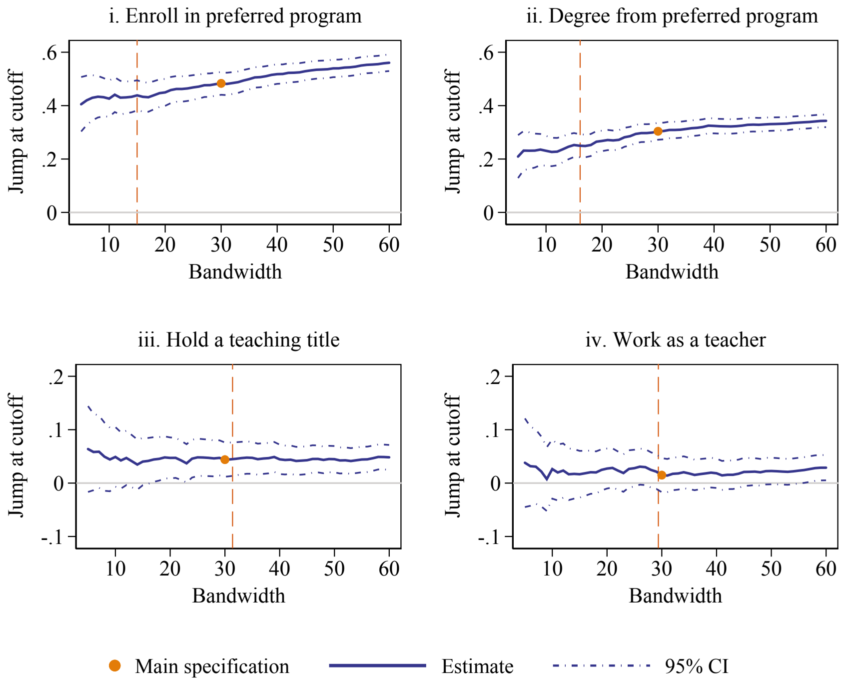


Panel J. First teaching job in a school located outside home province.



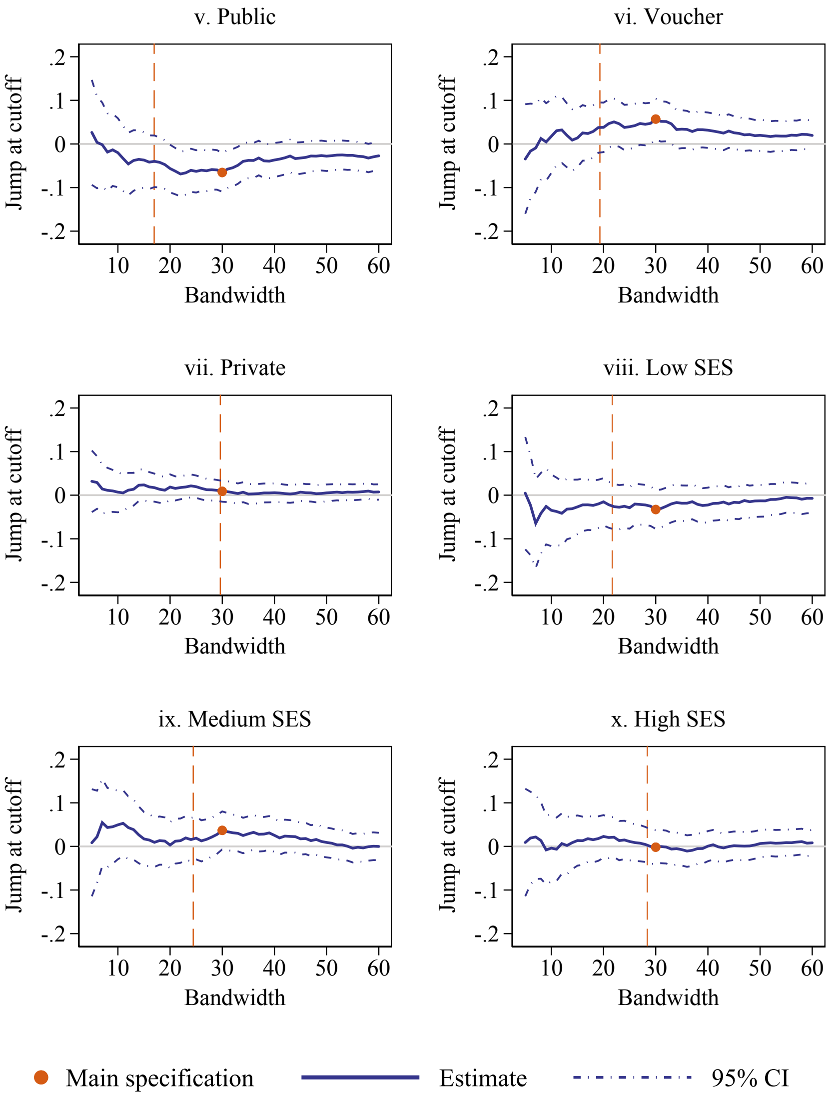
Notes: Figures on the left-hand side of each panel are for the full sample of applicants, while figures on the right-hand side are for the preferred sample of applicants who list a teaching program as both their top and next-best preference. Each blue dot represents the average outcome within a bin of size three. The vertical line indicates the admission cutoff to the preferred teaching program.

Figure A.11. Sensitivity of the first-stage and reduced-form estimates for the preferred sample at different bandwidths.



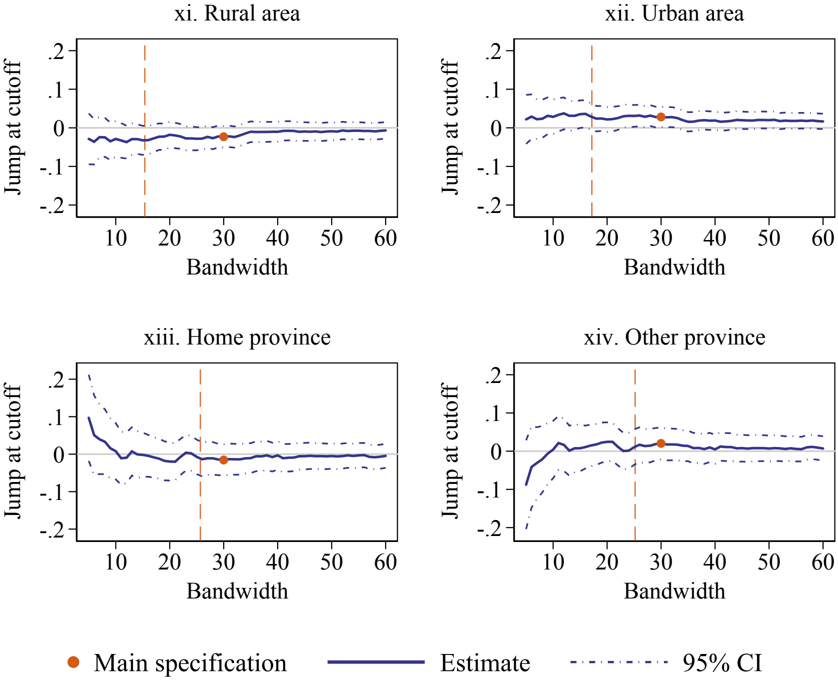
(Continued on next page.)

Figure A.11 (continued). Sensitivity to different bandwidths.



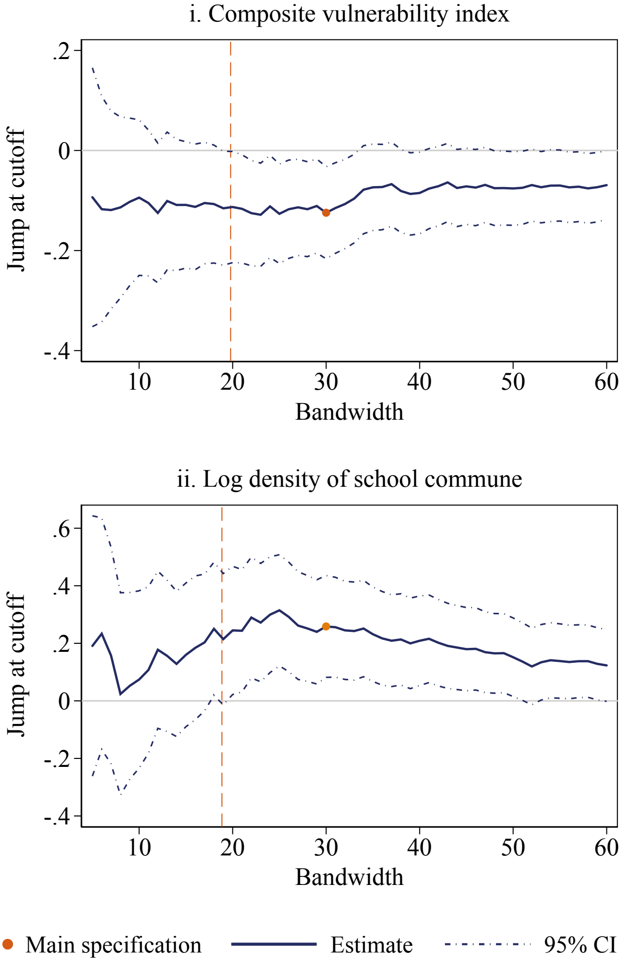
(Continued on next page.)

Figure A.11 (continued). Sensitivity to different bandwidths.



Notes: The above figures examine the sensitivity of the main first-stage and reduced-form results for the preferred sample. For each outcome, the solid blue line plots the estimated jump at the admission cutoff obtained using local linear regression with bandwidths ranging between five and 60. The 95% confidence intervals for the estimates are illustrated with dashed blue lines. The vertical orange line marks the optimal bandwidth according to the CCT selection procedure. As in the tables, all regressions include program-by-year and preference-type fixed effects, as well as a flexible control function in which the slope can vary above and below the cutoff by preference type.

Figure A.12. Additional bandwidth graphs for robustness checks.



Notes: The above figures show the reduced-form estimates at different bandwidths from five to 60 for two outcomes used for robustness checks. In the top panel, the outcome is a composite vulnerability index constructed by Principal Component Analysis. In the bottom panel, the outcome is the log population density of the commune where the teachers' school is located.

Table A.1. Specializations in pedagogy and share of titles granted in each sub-field 2007-2016.

Pedagogía en...	**Pedagogy in...**	**Percent**
Artes y Música	Art & Music	2.54
Ciencias	Natural Sciences	3.58
Educación Básica	Basic Education (Grades 1-8)	27.51
Educación Diferencial	Special Education	10.88
Educación Física	Physical Education	15.26
Educación de Párvulos	Early Education	16.46
Filosofía y Religión	Philosophy & Religion	1.49
Historia, Geografía y Ciencias Sociales	History, Geography & Social Sciences	5.65
Idiomas	Foreign Language	9.07
Lenguaje, Comunicación y/o Castellano	Language, Communication, and/or Spanish	4.18
Matemáticas y Computación	Mathematics & Computing	3.38

Notes: This table lists all possible specializations in pedagogy. The last column shows the share of all teaching titles granted between 2007 and 2016 that were earned within each particular specialization.

Table A.2. Descriptive statistics for applicants who apply to a teaching program in their top preference.

	Full sample		***Preferred sample***	
	Applicants	***Teachers***	***Applicants***	***Teachers***
	(1)	(2)	(3)	(4)
A. Personal background				
Female	.585	.663	.569	.642
Age at application	19.785	19.596	19.805	19.612
Resident of Santiago metro area	.265	.241	.284	.253
Log population density of home commune	5.737	5.544	5.778	5.574
Employed part- or full-time	.086	.073	.091	.078
B. Family background				
High household income	.332	.336	.325	.328
Household size	4.630	4.603	4.613	4.583
Number of parents alive	1.853	1.869	1.853	1.870
At least one parent works full-time	.717	.725	.717	.728
Father works in education	.070	.077	.069	.077
Mother works in education	.092	.099	.092	.098
Father's years of schooling	11.377	11.323	11.257	11.202
Mother's years of schooling	11.232	11.227	11.134	11.129
Neither parent went to college	.652	.660	.670	.679
C. Academic background				
Years since high school graduation	1.003	.904	1.019	.924
Graduate of public high school	.429	.419	.431	.423
Graduate of voucher high school	.527	.541	.535	.549
Graduate of private high school	.044	.040	.034	.028
High school grade point average	5.786	5.892	5.781	5.879
Score on college entrance exam	541.144	548.541	541.389	546.877
D. Application information				
Number of applications	4.792	4.847	4.874	4.962
Number of teaching programs	2.814	3.003	3.470	3.591
Apply only to teaching programs	.213	.235	.371	.375
Number of institutions	2.492	2.535	2.687	2.743
Apply only within home region	.538	.542	.479	.482
Receive any first-round offer	.680	.762	.669	.751
First-round offer for teaching	.539	.692	.589	.710
First-round offer for top program	.323	.449	.303	.398
Preference ranking of first offer	2.161	1.865	2.249	1.973
E. Outcomes				
Earn teaching title from top program	.227	.459	.216	.408
Earn any teaching title	.442	.883	.482	.894
Teach within 10 years of application	.419	1.000	.457	1.000
Years to first teaching contract	–	6.444	–	6.448
Observations	55,376	23,188	31,730	14,511

Notes: The full sample includes all applicants who list a teaching program as their top preference, while the preferred sample includes all applicants who list a teaching program as both a top and next-best preference. In columns 2 and 4, the sub-sample of teachers refers to applicants who work as a classroom teacher within 10 years of their first application to a teaching program.

Table A.3. Balance in baseline characteristics for the main samples of applicants and teachers.

	Full sample				**Preferred sample**			
	Applicants		***Teachers***		***Applicants***		***Teachers***	
	Mean (1)	Jump (2)	Mean (3)	Jump (4)	Mean (5)	Jump (6)	Mean (7)	Jump (8)
A. Personal background								
Female	.603	.007	.603	.007	.588	.018	.588	.018
	(.489)	(.011)	(.489)	(.011)	(.492)	(.015)	(.492)	(.015)
Age at application	19.764	-.006	19.764	-.006	19.773	-.002	19.773	-.002
	(2.259)	(.054)	(2.259)	(.054)	(2.305)	(.079)	(2.305)	(.079)
Resident of Santiago metro area	.292	.005	.292	.005	.310	.000	.310	.000
	(.455)	(.007)	(.455)	(.007)	(.463)	(.010)	(.463)	(.010)
Log population density of home commune	5.832	.007	5.832	.007	5.861	.013	5.861	.013
	(2.312)	(.045)	(2.312)	(.045)	(2.331)	(.063)	(2.331)	(.063)
Employed part- or full-time	.083	-.004	.083	-.004	.085	-.006	.085	-.006
	(.276)	(.007)	(.276)	(.007)	(.279)	(.010)	(.279)	(.010)
B. Family background								
High household income	.329	-.003	.333	-.004	.326	.002	.330	.004
	(.470)	(.012)	(.471)	(.017)	(.469)	(.016)	(.470)	(.022)
Household size	4.625	-.029	4.585	-.065	4.607	-.045	4.570	-.069
	(1.446)	(.034)	(1.386)	(.050)	(1.436)	(.046)	(1.379)	(.064)
Number of parents alive	1.852	.007	1.864	.015	1.854	.007	1.870	.007
	(.382)	(.009)	(.369)	(.014)	(.379)	(.013)	(.364)	(.017)
At least one parent works full-time	.721	.006	.734	.006	.726	.012	.738	.016
	(.448)	(.011)	(.442)	(.017)	(.446)	(.015)	(.440)	(.022)
Father works in education	.067	-.003	.075	-.002	.066	.004	.073	.002
	(.251)	(.007)	(.263)	(.011)	(.249)	(.010)	(.261)	(.014)
Mother works in education	.089	.002	.092	.008	.091	.006	.093	.018
	(.284)	(.007)	(.290)	(.011)	(.287)	(.010)	(.291)	(.015)

Continued on next page.

Table A.3 (continued). Balance in baseline characteristics for the main samples of applicants and teachers.

	Full sample				***Preferred sample***			
	Applicants		***Teachers***		***Applicants***		***Teachers***	
	Mean (1)	Jump (2)	Mean (3)	Jump (4)	Mean (5)	Jump (6)	Mean (7)	Jump (8)
Father's years of schooling	11.387	-.063	11.333	-.043	11.311	-.077	11.273	.027
	(3.744)	(.098)	(3.713)	(.144)	(3.713)	(.132)	(3.697)	(.185)
Mother's years of schooling	11.240	.028	11.236	.142	11.193	.040	11.203	.183
	(3.557)	(.090)	(3.527)	(.137)	(3.541)	(.125)	(3.538)	(.176)
Neither parent went to college	.653	.006	.664	-.003	.663	.011	.672	.003
	(.476)	(.012)	(.472)	(.018)	(.473)	(.016)	(.469)	(.023)
C. Academic background								
Years since high school graduation	1.002	.015	.877	.084	1.009	.044	.868	.108
	(1.957)	(.047)	(1.741)	(.066)	(1.968)	(.068)	(1.756)	(.089)
Graduate of public high school	.431	.011	.410	.005	.428	-.003	.408	-.014
	(.495)	(.012)	(.492)	(.018)	(.495)	(.015)	(.492)	(.023)
Graduate of voucher high school	.530	-.010	.558	.000	.540	.009	.566	.021
	(.499)	(.012)	(.497)	(.018)	(.498)	(.015)	(.496)	(.023)
Graduate of private high school	.039	.000	.032	-.005	.032	-.006	.026	-.007
	(.194)	(.005)	(.175)	(.006)	(.175)	(.006)	(.160)	(.008)
High school grade point average	5.787	.020	5.842	.013	5.796	.013	5.853	.001
	(.369)	(.008)**	(.361)	(.012)	(.372)	(.011)	(.363)	(.015)
p-value for joint significance test	.265		.473		.621		.495	
Number of individuals	25,709		11,635		14,718		7,348	

Notes: Each row reports the estimated jump at the cutoff for a different background variable. Estimates are obtained using local linear regressions with a bandwidth of 30, a rectangular kernel, program-by-year and preference-type fixed effects, as well as a flexible control function allowing the slope of the running variable to vary above and below the cutoff by preference type. Standard errors are clustered at the program-by-year level and are shown in parentheses. Stars denote significance levels: *** for $p < 0.01$; ** for $p < 0.05$; * for $p < 0.10$.

Table A.4. First-stage results for applicants who enroll in teaching programs.

	Full sample		***Preferred sample***	
	Mean (1)	Jump (2)	Mean (3)	Jump (4)
A. Institutional characteristics				
University	.996 (.064)	.002 (.002)	.997 (.055)	.002 (.002)
Traditional university	.727 (.445)	.065 (.012)***	.751 (.432)	.058 (.014)***
Accredited institution	.974 (.158)	.002 (.004)	.977 (.148)	.005 (.005)
Located in home region	.790 (.407)	.018 (.015)	.760 (.427)	.030 (.020)
Average achievement test score	542.852 (40.911)	14.697 (1.019)***	542.788 (38.570)	16.225 (1.277)***
Average high school grades	5.844 (.205)	.069 (.005)***	5.842 (.197)	.076 (.006)***
Share from public high school	.362 (.118)	-.012 (.002)***	.361 (.115)	-.016 (.003)***
Share from voucher high school	.539 (.104)	-.014 (.002)***	.544 (.100)	-.012 (.003)***
Share from private high school	.098 (.108)	.026 (.003)***	.094 (.100)	.029 (.004)***
B. Program characteristics				
Accredited program	.622 (.485)	.068 (.012)***	.638 (.481)	.077 (.016)***
Duration of study (years)	4.646 (.407)	.025 (.010)**	4.655 (.397)	.036 (.013)***
Average achievement test score	528.019 (43.139)	14.167 (.912)***	529.810 (41.445)	16.019 (1.115)***
Average high school grades	5.797 (.236)	.075 (.005)***	5.801 (.229)	.084 (.007)***
Share from public high school	.403 (.137)	-.008 (.003)***	.400 (.134)	-.011 (.003)***
Share from voucher high school	.547 (.120)	-.008 (.003)***	.552 (.117)	-.009 (.003)***
Share from private high school	.050 (.073)	.016 (.002)***	.048 (.067)	.020 (.003)***
Observations	10,821		6,879	

Notes: This table reports the jump in characteristics at the threshold for selected samples of applicants who enroll in a teaching program. Estimates are obtained via local linear regressions with a bandwidth of 30, rectangular kernel, program-by-year and preference-type fixed effects, as well as a flexible control function in which the slope of the running variable is allowed to vary above and below the cutoff for each preference type. Standard errors are clustered at the program-by-year level and are shown in parentheses. Significance levels are denoted with stars: *** for p-value < 0.01; ** for $p < 0.05$; * for $p < 0.10$.

Table A.5. Jumps in enrollment and completion of undergraduate programs.

	Full sample		***Preferred sample***	
	Mean (1)	Jump (2)	Mean (3)	Jump (4)
A. Preferred teaching program				
Immediately enrolled	.146	.500	.150	.515
	(.353)	(.019)***	(.357)	(.022)***
Ever enrolled	.199	.465	.195	.483
	(.399)	(.018)***	(.396)	(.022)***
Earned degree	.115	.290	.111	.304
	(.319)	(.013)***	(.315)	(.016)***
B. Preferred teaching field				
Immediately enrolled	.423	.293	.515	.213
	(.494)	(.017)***	(.500)	(.021)***
Ever enrolled	.516	.256	.604	.187
	(.500)	(.016)***	(.489)	(.019)***
Earned degree	.305	.162	.355	.128
	(.460)	(.013)***	(.478)	(.017)***
C. Any teaching program				
Immediately enrolled	.528	.199	.663	.064
	(.499)	(.015)***	(.473)	(.018)***
Ever enrolled	.639	.167	.766	.058
	(.480)	(.014)***	(.423)	(.016)***
Earned degree	.399	.112	.481	.044
	(.490)	(.012)***	(.500)	(.016)***
D. Any undergraduate program				
Immediately enrolled	.817	.045	.824	.024
	(.387)	(.011)***	(.381)	(.014)*
Ever enrolled	.971	.009	.973	.009
	(.167)	(.005)**	(.163)	(.006)
Earned degree	.686	.024	.693	.030
	(.464)	(.011)**	(.461)	(.015)**
Observations (enrollment outcomes)	14,553		8,355	
Observations (degree outcomes)	25,709		14,718	

Notes: Due to data limitations, the "ever enrolled" outcome and "immediately enrolled" outcome (i.e., enrollment in the year of first application) include application cohorts 2007-2010, whereas the completion outcomes also include cohorts 2004-2006. All estimates are obtained using local linear regressions with a rectangular kernel and a bandwidth of 30. The regression specifications include program-by-year and preference-type fixed effects, as well as a flexible control function that allows the slope of the running variable to vary above and below the cutoff for each different preference type. Standard errors are clustered at the program-by-year level and are shown in parentheses. Significance levels are denoted with stars: *** for p-value < 0.01; ** for $p < 0.05$; * for $p < 0.10$.

Table A.6. Specification checks for the preferred sample.

	(1)	(2)	(3)	(4)	(5)	(6)
A. Selection into teaching						
Hold a teaching title	.043	.044	.044	.042	.043	.037
	(.016)***	(.016)***	(.016)***	(.016)***	(.018)**	(.012)***
Work as a teacher	.009	.015	.015	.013	.021	.018
	(.016)	(.016)	(.016)	(.016)	(.017)	(.013)
B. Sorting by school sector						
Work in a public school	-.059	-.065	-.065	-.068	-.058	-.037
	(.022)***	(.023)***	(.023)***	(.022)***	(.024)**	(.018)**
Work in a voucher school	.057	.056	.057	.059	.041	.035
	(.023)**	(.024)**	(.024)**	(.024)**	(.026)	(.019)*
Work in a private school	.004	.009	.009	.009	.017	.003
	(.012)	(.012)	(.012)	(.012)	(.013)	(.010)
C. Sorting by student background						
Work in a low SES school	-.028	-.033	-.033	-.033	-.028	-.004
	(.022)	(.023)	(.023)	(.023)	(.024)	(.018)
Work in a medium SES school	.024	.037	.037	.037	.020	.003
	(.021)	(.022)*	(.022)*	(.022)*	(.023)	(.017)
Work in a high SES school	.004	-.002	-.002	-.001	.014	.001
	(.019)	(.020)	(.020)	(.020)	(.022)	(.016)
D. Sorting by region						
Work in a rural school	-.026	-.022	-.023	-.023	-.025	-.013
	(.014)*	(.014)	(.014)	(.014)*	(.015)*	(.012)
Work in an urban school	.032	.028	.028	.029	.028	.026
	(.013)**	(.014)**	(.014)**	(.014)**	(.014)**	(.011)**
Work in home province	-.026	-.016	-.016	-.017	-.009	-.003
	(.020)	(.021)	(.021)	(.021)	(.023)	(.017)
Work in other province	.030	.021	.020	.021	.014	.006
	(.020)	(.021)	(.021)	(.021)	(.023)	(.017)
Regression specification						
Local linear regression	X	X	X	X	X	—
Polynomial order	—	—	—	—	—	2nd
Program-by-year fixed effects	—	X	X	X	X	X
Preference type fixed effects	—	—	X	X	X	X
Preference type × control function	—	—	X	X	X	X
Additional covariates	—	—	—	X	X	—
Rectangular kernel	X	X	X	X	—	X
Triangular kernel	—	—	—	—	X	—
Observations (panel A)	14,718	14,718	14,718	14,718	14,718	31,730
Observations (panel B-D)	7,348	7,348	7,348	7,348	7,348	14,511

Notes: The baseline estimates are reported in column 3 for ease of comparison. The additional covariates in column 4 include age, gender, an indicator for high household income, and dummies for home region. Standard errors are clustered at the program-by-year level and shown in parentheses. Significance levels are denoted with stars: *** for p-value < 0.01; ** for $p < 0.05$; * for $p < 0.10$

Table A.7. Sensitivity of the preferred estimates to different sample restrictions.

	Main est.	***Excl. obs w/in x pts of cutoff*** $x=0$	$x=1$	$x=2$	***Excl. re-takes***
	(1)	(2)	(3)	(4)	(5)
A. Selection into teaching					
Hold a teaching title	.044	.046	.038	.038	.036
	(.016)***	(.017)***	(.018)**	(.019)*	(.021)*
Work as a teacher	.015	.015	.015	.013	.012
	(.016)	(.017)	(.019)	(.020)	(.021)
B. Sorting by school sector					
Work in a public school	-.065	-.065	-.064	-.085	-.078
	(.023)***	(.023)***	(.025)**	(.027)***	(.028)***
Work in a voucher school	.057	.054	.047	.072	.072
	(.024)**	(.025)**	(.026)*	(.028)***	(.029)**
Work in a private school	.009	.009	.013	.011	.016
	(.012)	(.012)	(.014)	(.015)	(.015)
C. Sorting by student background					
Work in a low SES school	-.033	-.038	-.042	-.060	-.040
	(.023)	(.023)*	(.026)	(.027)**	(.029)
Work in a medium SES school	.037	.043	.049	.073	.038
	(.022)*	(.023)*	(.025)*	(.027)***	(.029)
Work in a high SES school	-.002	-.003	-.011	-.016	.018
	(.020)	(.020)	(.021)	(.023)	(.025)
D. Sorting by region					
Work in a rural school	-.023	-.024	-.026	-.023	-.026
	(.014)	(.014)*	(.016)*	(.016)	(.017)
Work in an urban school	.028	.028	.029	.029	.033
	(.014)**	(.014)**	(.015)*	(.016)*	(.016)**
Work in home province	-.016	-.008	-.013	-.019	-.019
	(.021)	(.022)	(.023)	(.025)	(.027)
Work in other province	.020	.013	.017	.022	.023
	(.021)	(.022)	(.024)	(.025)	(.027)
Observations (panel A)	14,718	14,547	14,033	13,536	9,833
Observations (panel B-D)	7,348	7,257	6,991	6,724	4,917

Notes: Column 1 replicates the baseline results to facilitate comparison with estimates in the other columns. Columns 2 through 4 perform donut RD, excluding points right on the cutoff, within one point of the cutoff, and within two points of the cutoff respectively. In column 5, the sample is restricted to people who submitted their first application in the same year that they first wrote the entrance exam. All estimates are obtained using local linear regression with a rectangular kernel and a bandwidth of 30. Standard errors are clustered at the program-by-year level and are shown in parentheses. Significance levels are denoted with stars: *** for p-value < 0.01; ** for $p < 0.05$; * for $p < 0.10$.

Table A.8. Sensitivity of the preferred estimates to definition of the outcomes.

	Any workplace		***Primary workplace***	
	Mean	RF	Mean	RF
	(1)	(2)	(3)	(4)
A. Selection into teaching				
Work as a teacher	.462	.015	.462	.015
	(.499)	(.016)	(.499)	(.016)
B. Sorting by school sector				
Work in a public school	.364	-.065	.355	-.062
	(.481)	(.023)***	(.479)	(.023)***
Work in a voucher school	.597	.057	.586	.056
	(.491)	(.024)**	(.493)	(.024)**
Work in a private school	.061	.009	.058	.005
	(.239)	(.012)	(.234)	(.012)
C. Sorting by student background				
Work in a low SES school	.462	-.033	.451	-.036
	(.499)	(.023)	(.498)	(.023)
Work in a medium SES school	.333	.037	.320	.040
	(.471)	(.022)*	(.466)	(.022)*
Work in a high SES school	.236	-.002	.229	-.004
	(.425)	(.020)	(.420)	(.019)
D. Sorting by region				
Work in a rural school	.106	-.023	.100	-.025
	(.308)	(.014)	(.300)	(.014)*
Work in an urban school	.903	.028	.900	.025
	(.296)	(.014)**	(.300)	(.014)*
Work in home province	.704	-.016	.701	-.023
	(.456)	(.021)	(.458)	(.021)
Work in other province	.302	.020	.299	.023
	(.459)	(.021)	(.458)	(.021)
Observations (panel A)	14,718		14,718	
Observations (panel B-D)	7,348		7,348	

Notes: The baseline results are replicated here in column 2 to facilitate comparison with estimates in column 4. In column 4, the characteristics of the primary workplace (i.e., where the individual is contracted to work the most hours) are used to define the probabilities. All estimates are obtained using local linear regression with a rectangular kernel and a bandwidth of 30. The regressions include program-by-year fixed effects and preference-type fixed effects. Standard errors are clustered at the program-by-year level and are shown in parentheses. Significance levels are denoted with stars: *** for p-value < 0.01; ** for $p < 0.05$; * for $p < 0.10$.

Table A.9. Composite measure of school vulnerability.

Component	***Score coefficient***
Hours worked in public school	0.601
Hours worked in low SES school	0.629
Hours worked in rural school	0.487
Eigenvalue	1.812
Share of explained variation	0.604
Cronbach's alpha (reliability)	0.667

Notes: This table reports the score coefficients used to construct the school vulnerability index.

Table A.10. Results for composite index and continuous measure of urbanicity.

	Mean (1)	Jump (2)
Composite vulnerability index	-.020 (.998)	-.124 (.047)***
Log density of school commune	5.629 (2.336)	.259 (.090)***
Observations	7,348	

Notes: This table reports two different reduced-form estimates for the preferred sample. In the first row, the outcome is the composite vulnerability index constructed by Principal Component Analysis. In the second row, the outcome is the log population density of the commune where the school is located. Standard errors are clustered at the program-by-year level and are shown in parentheses. Significance levels are denoted with stars: *** for p-value < 0.01; ** for $p < 0.05$; * for $p < 0.10$.

Table A.11. Heterogeneity analysis for above versus below median quality programs.

	Above median quality			***Below median quality***		
	Mean	RF	SS	Mean	RF	SS
	(1)	(2)	(3)	(4)	(5)	(6)
A. Selection into teaching						
Hold a teaching title	.499	.043	—	.437	.047	—
	(.500)	(.020)**		(.496)	(.027)*	
Work as a teacher	.473	.018	.050	.435	.013	.072
	(.499)	(.020)	(.054)	(.496)	(.028)	(.149)
B. Sorting by school sector						
Work in a public school	.333	-.076	-.143	.449	-.039	-.157
	(.471)	(.027)***	(.051)***	(.498)	(.042)	(.166)
Work in a voucher school	.618	.063	.117	.540	.043	.171
	(.486)	(.028)**	(.053)**	(.499)	(.044)	(.175)
Work in a private school	.072	.018	.034	.031	-.008	-.033
	(.258)	(.016)	(.030)	(.172)	(.016)	(.063)
C. Sorting by student background						
Work in a low SES school	.421	-.037	-.069	.574	-.021	-.085
	(.494)	(.028)	(.052)	(.495)	(.039)	(.158)
Work in a medium SES school	.347	.036	.067	.295	.036	.146
	(.476)	(.028)	(.053)	(.456)	(.036)	(.144)
Work in a high SES school	.263	.002	.004	.162	-.009	-.035
	(.440)	(.025)	(.047)	(.369)	(.030)	(.122)

Continued on next page.

Table A.11 (continued). Heterogeneity analysis for above versus below median quality programs.

	Above median quality			***Below median quality***		
	Mean	RF	SS	Mean	RF	SS
	(1)	(2)	(3)	(4)	(5)	(6)
D. Sorting by region						
Work in a rural school	.085	-.029	-.054	.166	-.007	-.027
	(.278)	(.015)*	(.028)*	(.372)	(.031)	(.126)
Work in an urban school	.925	.027	.051	.844	.028	.112
	(.264)	(.015)*	(.027)*	(.363)	(.029)	(.118)
Work in home province	.698	-.028	-.052	.722	.016	.066
	(.459)	(.026)	(.048)	(.448)	(.038)	(.154)
Work in other province	.309	.034	.064	.283	-.014	-.057
	(.462)	(.026)	(.048)	(.451)	(.038)	(.155)
First-stage F-stat (panel A)		620.796			59.615	
First-stage F-stat (panel B-D)		770.732			57.934	
Observations (panel A)		10,078			4,460	
Observations (panel B-D)		5,100			2,248	

Notes: This table reports the main RD results for two sub-samples of the preferred sample: applicants whose top teaching program is above- versus below-median quality in their admission year. Each row focuses on a different labor market outcome. Columns labeled "RF" report reduced-form estimates of the effect of getting admitted to the most-preferred teaching program. Columns labeled "SS" report second-stage estimates in which completing the most-preferred teaching program is instrumented with distance to the admission cutoff for that program. All estimates are obtained using local linear regressions with a rectangular kernel and a bandwidth of 30. The regression specifications include program-by-year and preference-type fixed effects, as well as a flexible control function in which the slope of the running variable is allowed to vary above and below the cutoff by preference type. Robust standard errors are shown in parentheses. Significance levels are denoted with stars: *** for p-value < 0.01; ** for $p < 0.05$; * for $p < 0.10$.

Table A.12. Heterogeneity analysis by type of high school attended.

	Graduate of public high school			***Graduate of other high school***		
	Mean (1)	RF (2)	SS (3)	Mean (4)	RF (5)	SS (6)
A. Selection into teaching						
Hold a teaching title	.453	.023	—	.502	.047	—
	(.498)	(.025)		(.500)	(.022)**	
Work as a teacher	.439	-.003	-.011	.479	.020	.065
	(.496)	(.025)	(.086)	(.500)	(.022)	(.069)
B. Sorting by school sector						
Work in a public school	.447	-.037	-.080	.307	-.052	-.110
	(.497)	(.040)	(.083)	(.461)	(.030)*	(.062)*
Work in a voucher school	.525	.007	.015	.647	.065	.139
	(.500)	(.041)	(.084)	(.478)	(.032)**	(.066)**
Work in a private school	.048	.031	.067	.069	-.006	-.013
	(.215)	(.018)*	(.037)*	(.254)	(.018)	(.037)
C. Sorting by student background						
Work in a low SES school	.530	-.039	-.084	.415	-.029	-.063
	(.499)	(.039)	(.082)	(.493)	(.032)	(.066)
Work in a medium SES school	.300	-.005	-.010	.356	.061	.130
	(.458)	(.037)	(.077)	(.479)	(.032)*	(.067)*
Work in a high SES school	.196	.044	.093	.263	-.024	-.051
	(.397)	(.032)	(.066)	(.441)	(.030)	(.061)

Continued on next page.

Table A.12 (continued). Heterogeneity analysis by type of high school attended.

	Graduate of public high school			***Graduate of other high school***		
	Mean (1)	RF (2)	SS (3)	Mean (4)	RF (5)	SS (6)
D. Sorting by region						
Work in a rural school	.135	-.015	-.033	.087	-.019	-.040
	(.342)	(.027)	(.056)	(.281)	(.019)	(.039)
Work in an urban school	.878	.024	.051	.921	.027	.057
	(.327)	(.026)	(.053)	(.270)	(.018)	(.037)
Work in home province	.694	.015	.032	.711	-.027	-.057
	(.461)	(.037)	(.078)	(.453)	(.030)	(.063)
Work in other province	.310	-.006	-.012	.296	.029	.061
	(.463)	(.037)	(.078)	(.457)	(.031)	(.063)
First-stage F-stat (panel A)		428.126			351.336	
First-stage F-stat (panel B-D)		378.249			442.488	
Observations (panel A)		6,375			8,352	
Observations (panel B-D)		3,077			4,324	

Notes: This table reports the main RD results for two subsamples of the preferred sample: applicants who graduated from a public high school versus a voucher or private high school. Each row focuses on a different labor market outcome. All estimates are obtained using local linear regressions with a rectangular kernel and a bandwidth of 30. The regression specifications include program-by-year and preference-type fixed effects, as well as a flexible control function in which the slope of the running variable is allowed to vary above and below the cutoff for each different preference type. Standard errors are clustered at the program-by-year level and are shown in parentheses. Significance levels are denoted with stars: *** for p-value < 0.01; ** for $p < 0.05$; * for $p < 0.10$.

References

Altonji, J. G. and Pierret, C. R. (2001). Employer Learning and Statistical Discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.

Ávalos, B. and Valenzuela, J. P. (2016). Education for all and attrition/retention of new teachers: A trajectory study in Chile. *International Journal of Educational Development*, 49:279–290.

Baker, B. D. and Dickerson, J. L. (2006). Charter Schools, Teacher Labor Market Deregulation, and Teacher Quality: Evidence From the Schools and Staffing Survey. *Educational Policy*, 20(5):752–778.

Ballou, D. (1996). Do Public Schools Hire the Best Applicants? *The Quarterly Journal of Economics*, 111(1):97–133.

Bonesrønning, H., Falch, T., and Strøm, B. (2005). Teacher sorting, teacher quality, and student composition. *European Economic Review*, 49(2):457–483.

Boyd, D., Lankford, H., Loeb, S., Ronfeldt, M., and Wyckoff, J. (2011). The role of teacher quality in retention and hiring: Using applications to transfer to uncover preferences of teachers and schools. *Journal of Policy Analysis and Management*, 30(1):88–110.

Boyd, D., Lankford, H., Loeb, S., and Wyckoff, J. (2005). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management*, 24(1):113–132.

Burian-Fitzgerald, M. and Harris, D. (2004). Teacher Recruitment and Teacher Quality? Are Charter Schools Different? Policy Report Number 20.

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

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, 104(9):2633–2679.

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review*, 24(4):377–392.

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2010). Teacher Credentials and Student Achievement in High School A Cross-Subject Analysis with Student Fixed Effects. *Journal of Human Resources*, 45(3):655–681.

Cox, C. (2010). Educational Inequality in Latin America: Patterns, Policies and Issues. In *Educational Inequality Around the World*, pages 33–58.

Ehrenberg, R. G. and Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from High School and Beyond. *Economics of Education Review*, 13(1):1–17.

Falch, T. and Strøm, B. (2005). Teacher turnover and non-pecuniary factors. *Economics of Education Review*, 24(6):611–631.

Figlio, D. N. (1997). Teacher salaries and teacher quality. *Economics Letters*, 55(2):267–271.

Gelman, A. and Imbens, G. (2019). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Goldhaber, D., Choi, H.-J., and Cramer, L. (2007). A descriptive analysis of the distribution of NBPTS-certified teachers in North Carolina. *Economics of Education Review*, 26(2):160–172.

Goldhaber, D., Lavery, L., and Theobald, R. (2015). Uneven Playing Field? Assessing the Teacher Quality Gap Between Advantaged and Disadvantaged Students. *Educational Researcher*, 44(5):293–307.

Hanushek, E. A., Kain, J. F., and Rivkin, S. G. (2004). Why Public Schools Lose Teachers. *The Journal of Human Resources*, 39(2):326–354.

Hinrichs, P. (2014). What Kind of Teachers Are Schools Looking For? Evidence from a Randomized Field Experiment. *Forthcoming in Journal of Economic Behavior Organization*.

Jackson, C. (2009). Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation. *Journal of Labor Economics*, 27(2):213–256.

Kalogrides, D., Loeb, S., and Béteille, T. (2013). Systematic Sorting: Teacher Characteristics and Class Assignments. *Sociology of Education*, 86(2):103–123.

Lankford, H., Loeb, S., and Wyckoff, J. (2002). Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis. *Educational Evaluation and Policy Analysis*, 24(1):37–62.

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

Loeb, S. and Reininger, M. (2004). Public policy and teacher labor markets: What we know and why it matters. *The Education Policy Center at Michigan State University*.

Luschei, T. F. and Jeong, D. W. (2018). Is Teacher Sorting a Global Phenomenon? Cross-National Evidence on the Nature and Correlates of Teacher Quality Opportunity Gaps. *Educational Researcher*, 47(9):556–576.

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

Ministerio de Educación (2018). Mi Futuro: Estadísticas por carrera [My Future: Statistics by career]. url: https://www.mifuturo.cl.

Monk, D. H. (2007). Recruiting and Retaining High-Quality Teachers in Rural Areas. *The Future of Children*, 17(1):155–174.

OECD (2013). *PISA 2012 Results: What Makes Schools Successful (Volume IV): Resources, Policies and Practices*. OECD Publishing, Paris.

Papay, J. P. and Kraft, M. A. (2016). The Productivity Costs of Inefficient Hiring Practices: Evidence From Late Teacher Hiring. *Journal of Policy Analysis and Management*, 35(4):791–817.

Rice, J. K. (2003). *Teacher Quality: Understanding the Effectiveness of Teacher Attributes*. Economic Policy Institute.

Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73(2):417–458.

Rockoff, J. E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *The American Economic Review*, 94(2):247–252.

Santiago, P., Benavides, F., Danielson, C., Goe, L., and Nusche, D. (2013). *Teacher Evaluation in Chile 2013*. OECD Reviews of Evaluation and Assessment in Education. OECD Publishing, Paris.

Santiago, P., Fiszbein, A., García Jaramillo, S., and Radinger, T. (2017). *OECD Reviews of School Resources: Chile 2017*. OECD Reviews of School Resources. OECD Publishing, Paris.

Sass, T. R., Hannaway, J., Xu, Z., Figlio, D. N., and Feng, L. (2012). Value added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics*, 72(2):104–122.

Scafidi, B., Sjoquist, D. L., and Stinebrickner, T. R. (2007). Race, poverty, and teacher mobility. *Economics of Education Review*, 26(2):145–159.

Steele, J. L., Pepper, M. J., Springer, M. G., and Lockwood, J. R. (2015). The distribution and mobility of effective teachers: Evidence from a large, urban school district. *Economics of Education Review*, 48:86–101.

Thiemann, P. (2019). Inequality in Education Outcomes: The Role of Sorting among Students, Teachers, and Schools. Working Paper.

Economic Studies

1987:1 Haraldson, Marty. To Care and To Cure. A linear programming approach to national health planning in developing countries. 98 pp.

1989:1 Chryssanthou, Nikos. The Portfolio Demand for the ECU. A Transaction Cost Approach. 42 pp.

1989:2 Hansson, Bengt. Construction of Swedish Capital Stocks, 1963-87. An Application of the Hulten-Wykoff Studies. 37 pp.

1989:3 Choe, Byung-Tae. Some Notes on Utility Functions Demand and Aggregation. 39 pp.

1989:4 Skedinger, Per. Studies of Wage and Employment Determination in the Swedish Wood Industry. 89 pp.

1990:1 Gustafson, Claes-Håkan. Inventory Investment in Manufacturing Firms. Theory and Evidence. 98 pp.

1990:2 Bantekas, Apostolos. The Demand for Male and Female Workers in Swedish Manufacturing. 56 pp.

1991:1 Lundholm, Michael. Compulsory Social Insurance. A Critical Review. 109 pp.

1992:1 Sundberg, Gun. The Demand for Health and Medical Care in Sweden. 58 pp.

1992:2 Gustavsson, Thomas. No Arbitrage Pricing and the Term Structure of Interest Rates. 47 pp.

1992:3 Elvander, Nils. Labour Market Relations in Sweden and Great Britain. A Comparative Study of Local Wage Formation in the Private Sector during the 1980s. 43 pp.

12 Dillén, Mats. Studies in Optimal Taxation, Stabilization, and Imperfect Competition. 1993. 143 pp.

13 Banks, Ferdinand E.. A Modern Introduction to International Money, Banking and Finance. 1993. 303 pp.

14 Mellander, Erik. Measuring Productivity and Inefficiency Without Quantitative Output Data. 1993. 140 pp.

15 Ackum Agell. Susanne. Essays on Work and Pay. 1993. 116 pp.

16 Eriksson, Claes. Essays on Growth and Distribution. 1994. 129 pp.

17 Banks, Ferdinand E.. A Modern Introduction to International Money, Banking and Finance. 2nd version, 1994. 313 pp.

18 Apel, Mikael. Essays on Taxation and Economic Behavior. 1994. 144 pp.

19 Dillén, Hans. Asset Prices in Open Monetary Economies. A Contingent Claims Approach. 1994. 100 pp.

20 Jansson, Per. Essays on Empirical Macroeconomics. 1994. 146 pp.

21 Banks, Ferdinand E.. A Modern Introduction to International Money, Banking, and Finance. 3rd version, 1995. 313 pp.

22 Dufwenberg, Martin. On Rationality and Belief Formation in Games. 1995. 93 pp.

23 Lindén, Johan. Job Search and Wage Bargaining. 1995. 127 pp.

24 Shahnazarian, Hovick. Three Essays on Corporate Taxation. 1996. 112 pp.

25 Svensson, Roger. Foreign Activities of Swedish Multinational Corporations. 1996. 166 pp.

26 Sundberg, Gun. Essays on Health Economics. 1996. 174 pp.

27 Sacklén, Hans. Essays on Empirical Models of Labor Supply. 1996. 168 pp.

28 Fredriksson, Peter. Education, Migration and Active Labor Market Policy. 1997. 106 pp.

29 Ekman, Erik. Household and Corporate Behaviour under Uncertainty. 1997. 160 pp.

30 Stoltz, Bo. Essays on Portfolio Behavior and Asset Pricing. 1997. 122 pp.

31 Dahlberg, Matz. Essays on Estimation Methods and Local Public Economics. 1997. 179 pp.

32 Kolm, Ann-Sofie. Taxation, Wage Formation, Unemployment and Welfare. 1997. 162 pp.

33 Boije, Robert. Capitalisation, Efficiency and the Demand for Local Public Services. 1997. 148 pp.

34 Hort, Katinka. On Price Formation and Quantity Adjustment in Swedish Housing Markets. 1997. 185 pp.

35 Lindström, Thomas. Studies in Empirical Macroeconomics. 1998. 113 pp.

36 Hemström, Maria. Salary Determination in Professional Labour Markets. 1998. 127 pp.

37 Forsling, Gunnar. Utilization of Tax Allowances and Corporate Borrowing. 1998. 96 pp.

38 Nydahl, Stefan. Essays on Stock Prices and Exchange Rates. 1998. 133 pp.

39 Bergström, Pål. Essays on Labour Economics and Econometrics. 1998. 163 pp.

40 Heiborn, Marie. Essays on Demographic Factors and Housing Markets. 1998. 138 pp.

41 Åsberg, Per. Four Essays in Housing Economics. 1998. 166 pp.

42 Hokkanen, Jyry. Interpreting Budget Deficits and Productivity Fluctuations. 1998. 146 pp.

43 Lunander, Anders. Bids and Values. 1999. 127 pp.

44 Eklöf, Matias. Studies in Empirical Microeconomics. 1999. 213 pp.

45 Johansson, Eva. Essays on Local Public Finance and Intergovernmental Grants. 1999. 156 pp.

46 Lundin, Douglas. Studies in Empirical Public Economics. 1999. 97 pp.

47 Hansen, Sten. Essays on Finance, Taxation and Corporate Investment. 1999. 140 pp.

48 Widmalm, Frida. Studies in Growth and Household Allocation. 2000. 100 pp.

49 Arslanogullari, Sebastian. Household Adjustment to Unemployment. 2000. 153 pp.

50 Lindberg, Sara. Studies in Credit Constraints and Economic Behavior. 2000. 135 pp.

51 Nordblom, Katarina. Essays on Fiscal Policy, Growth, and the Importance of Family Altruism. 2000. 105 pp.

52 Andersson, Björn. Growth, Saving, and Demography. 2000. 99 pp.

53 Åslund, Olof. Health, Immigration, and Settlement Policies. 2000. 224 pp.

54 Bali Swain, Ranjula. Demand, Segmentation and Rationing in the Rural Credit Markets of Puri. 2001. 160 pp.

55 Löfqvist, Richard. Tax Avoidance, Dividend Signaling and Shareholder Taxation in an Open Economy. 2001. 145 pp.

56 Vejsiu, Altin. Essays on Labor Market Dynamics. 2001. 209 pp.

57 Zetterström, Erik. Residential Mobility and Tenure Choice in the Swedish Housing Market. 2001. 125 pp.

58 Grahn, Sofia. Topics in Cooperative Game Theory. 2001. 106 pp.

59 Laséen, Stefan. Macroeconomic Fluctuations and Microeconomic Adjustments. Wages, Capital, and Labor Market Policy. 2001. 142 pp.

60 Arnek, Magnus. Empirical Essays on Procurement and Regulation. 2002. 155 pp.

61 Jordahl, Henrik. Essays on Voting Behavior, Labor Market Policy, and Taxation. 2002. 172 pp.

62 Lindhe, Tobias. Corporate Tax Integration and the Cost of Capital. 2002. 102 pp.

63 Hallberg, Daniel. Essays on Household Behavior and Time-Use. 2002. 170 pp.

64 Larsson, Laura. Evaluating Social Programs: Active Labor Market Policies and Social Insurance. 2002. 126 pp.

65 Bergvall, Anders. Essays on Exchange Rates and Macroeconomic Stability. 2002. 122 pp.

66 Nordström Skans, Oskar. Labour Market Effects of Working Time Reductions and Demographic Changes. 2002. 118 pp.

67 Jansson, Joakim. Empirical Studies in Corporate Finance, Taxation and Investment. 2002. 132 pp.

68 Carlsson, Mikael. Macroeconomic Fluctuations and Firm Dynamics: Technology, Production and Capital Formation. 2002. 149 pp.

69 Eriksson, Stefan. The Persistence of Unemployment: Does Competition between Employed and Unemployed Job Applicants Matter? 2002. 154 pp.

70 Huitfeldt, Henrik. Labour Market Behaviour in a Transition Economy: The Czech Experience. 2003. 110 pp.

71 Johnsson, Richard. Transport Tax Policy Simulations and Satellite Accounting within a CGE Framework. 2003. 84 pp.

72 Öberg, Ann. Essays on Capital Income Taxation in the Corporate and Housing Sectors. 2003. 183 pp.

73 Andersson, Fredrik. Causes and Labor Market Consequences of Producer Heterogeneity. 2003. 197 pp.

74 Engström, Per. Optimal Taxation in Search Equilibrium. 2003. 127 pp.

75 Lundin, Magnus. The Dynamic Behavior of Prices and Investment: Financial Constraints and Customer Markets. 2003. 125 pp.

76 Ekström, Erika. Essays on Inequality and Education. 2003. 166 pp.

77 Barot, Bharat. Empirical Studies in Consumption, House Prices and the Accuracy of European Growth and Inflation Forecasts. 2003. 137 pp.

78 Österholm, Pär. Time Series and Macroeconomics: Studies in Demography and Monetary Policy. 2004. 116 pp.

79 Bruér, Mattias. Empirical Studies in Demography and Macroeconomics. 2004. 113 pp.

80 Gustavsson, Magnus. Empirical Essays on Earnings Inequality. 2004. 154 pp.

81 Toll, Stefan. Studies in Mortgage Pricing and Finance Theory. 2004. 100 pp.

82 Hesselius, Patrik. Sickness Absence and Labour Market Outcomes. 2004. 109 pp.

83 Häkkinen, Iida. Essays on School Resources, Academic Achievement and Student Employment. 2004. 123 pp.

84 Armelius, Hanna. Distributional Side Effects of Tax Policies: An Analysis of Tax Avoidance and Congestion Tolls. 2004. 96 pp.

85 Ahlin, Åsa. Compulsory Schooling in a Decentralized Setting: Studies of the Swedish Case. 2004. 148 pp.

86 Heldt, Tobias. Sustainable Nature Tourism and the Nature of Tourists' Cooperative Behavior: Recreation Conflicts, Conditional Cooperation and the Public Good Problem. 2005. 148 pp.

87 Holmberg, Pär. Modelling Bidding Behaviour in Electricity Auctions: Supply Function Equilibria with Uncertain Demand and Capacity Constraints. 2005. 43 pp.

88 Welz, Peter. Quantitative new Keynesian macroeconomics and monetary policy 2005. 128 pp.

89 Ågren, Hanna. Essays on Political Representation, Electoral Accountability and Strategic Interactions. 2005. 147 pp.

90 Budh, Erika. Essays on environmental economics. 2005. 115 pp.

91 Chen, Jie. Empirical Essays on Housing Allowances, Housing Wealth and Aggregate Consumption. 2005. 192 pp.

92 Angelov, Nikolay. Essays on Unit-Root Testing and on Discrete-Response Modelling of Firm Mergers. 2006. 127 pp.

93 Savvidou, Eleni. Technology, Human Capital and Labor Demand. 2006. 151 pp.

94 Lindvall, Lars. Public Expenditures and Youth Crime. 2006. 112 pp.

95 Söderström, Martin. Evaluating Institutional Changes in Education and Wage Policy. 2006. 131 pp.

96 Lagerström, Jonas. Discrimination, Sickness Absence, and Labor Market Policy. 2006. 105 pp.

97 Johansson, Kerstin. Empirical essays on labor-force participation, matching, and trade. 2006. 168 pp.

98 Ågren, Martin. Essays on Prospect Theory and the Statistical Modeling of Financial Returns. 2006. 105 pp.

99 Nahum, Ruth-Aïda. Studies on the Determinants and Effects of Health, Inequality and Labour Supply: Micro and Macro Evidence. 2006. 153 pp.

100 Žamac, Jovan. Education, Pensions, and Demography. 2007. 105 pp.

101 Post, Erik. Macroeconomic Uncertainty and Exchange Rate Policy. 2007. 129 pp.

102 Nordberg, Mikael. Allies Yet Rivals: Input Joint Ventures and Their Competitive Effects. 2007. 122 pp.

103 Johansson, Fredrik. Essays on Measurement Error and Nonresponse. 2007. 130 pp.

104 Haraldsson, Mattias. Essays on Transport Economics. 2007. 104 pp.

105 Edmark, Karin. Strategic Interactions among Swedish Local Governments. 2007. 141 pp.

106 Oreland, Carl. Family Control in Swedish Public Companies. Implications for Firm Performance, Dividends and CEO Cash Compensation. 2007. 121 pp.

107 Andersson, Christian. Teachers and Student Outcomes: Evidence using Swedish Data. 2007. 154 pp.

108 Kjellberg, David. Expectations, Uncertainty, and Monetary Policy. 2007. 132 pp.

109 Nykvist, Jenny. Self-employment Entry and Survival - Evidence from Sweden. 2008. 94 pp.

110 Selin, Håkan. Four Empirical Essays on Responses to Income Taxation. 2008. 133 pp.

111 Lindahl, Erica. Empirical studies of public policies within the primary school and the sickness insurance. 2008. 143 pp.

112 Liang, Che-Yuan. Essays in Political Economics and Public Finance. 2008. 125 pp.

113 Elinder, Mikael. Essays on Economic Voting, Cognitive Dissonance, and Trust. 2008. 120 pp.

114 Grönqvist, Hans. Essays in Labor and Demographic Economics. 2009. 120 pp.

115 Bengtsson, Niklas. Essays in Development and Labor Economics. 2009. 93 pp.

116 Vikström, Johan. Incentives and Norms in Social Insurance: Applications, Identification and Inference. 2009. 205 pp.

117 Liu, Qian. Essays on Labor Economics: Education, Employment, and Gender. 2009. 133 pp.

118 Glans, Erik. Pension reforms and retirement behaviour. 2009. 126 pp.

119 Douhan, Robin. Development, Education and Entrepreneurship. 2009.

120 Nilsson, Peter. Essays on Social Interactions and the Long-term Effects of Early-life Conditions. 2009. 180 pp.

121 Johansson, Elly-Ann. Essays on schooling, gender, and parental leave. 2010. 131 pp.

122 Hall, Caroline. Empirical Essays on Education and Social Insurance Policies. 2010. 147 pp.

123 Enström-Öst, Cecilia. Housing policy and family formation. 2010. 98 pp.

124 Winstrand, Jakob. Essays on Valuation of Environmental Attributes. 2010. 96 pp.

125 Söderberg, Johan. Price Setting, Inflation Dynamics, and Monetary Policy. 2010. 102 pp.

126 Rickne, Johanna. Essays in Development, Institutions and Gender. 2011. 138 pp.

127 Hensvik, Lena. The effects of markets, managers and peers on worker outcomes. 2011. 179 pp.

128 Lundqvist, Heléne. Empirical Essays in Political and Public. 2011. 157 pp.

129 Bastani, Spencer. Essays on the Economics of Income Taxation. 2012. 257 pp.

130 Corbo, Vesna. Monetary Policy, Trade Dynamics, and Labor Markets in Open Economies. 2012. 262 pp.

131 Nordin, Mattias. Information, Voting Behavior and Electoral Accountability. 2012. 187 pp.

132 Vikman, Ulrika. Benefits or Work? Social Programs and Labor Supply. 2013. 161 pp.

133 Ek, Susanne. Essays on unemployment insurance design. 2013. 136 pp.

134 Österholm, Göran. Essays on Managerial Compensation. 2013. 143 pp.

135 Adermon, Adrian. Essays on the transmission of human capital and the impact of technological change. 2013. 138 pp.

136 Kolsrud, Jonas. Insuring Against Unemployment 2013. 140 pp.

137 Hanspers, Kajsa. Essays on Welfare Dependency and the Privatization of Welfare Services. 2013. 208 pp.

138 Persson, Anna. Activation Programs, Benefit Take-Up, and Labor Market Attachment. 2013. 164 pp.

139 Engdahl, Mattias. International Mobility and the Labor Market. 2013. 216 pp.

140 Krzysztof Karbownik. Essays in education and family economics. 2013. 182 pp.

141 Oscar Erixson. Economic Decisions and Social Norms in Life and Death Situations. 2013. 183 pp.

142 Pia Fromlet. Essays on Inflation Targeting and Export Price Dynamics. 2013. 145 pp.

143 Daniel Avdic. Microeconometric Analyses of Individual Behavior in Public Welfare Systems. Applications in Health and Education Economics. 2014. 176 pp.

144 Arizo Karimi. Impacts of Policies, Peers and Parenthood on Labor Market Outcomes. 2014. 221 pp.

145 Karolina Stadin. Employment Dynamics. 2014. 134 pp.

146 Haishan Yu. Essays on Environmental and Energy Economics. 132 pp.

147 Martin Nilsson. Essays on Health Shocks and Social Insurance. 139 pp.

148 Tove Eliasson. Empirical Essays on Wage Setting and Immigrant Labor Market Opportunities. 2014. 144 pp.

149 Erik Spector. Financial Frictions and Firm Dynamics. 2014. 129 pp.

150 Michihito Ando. Essays on the Evaluation of Public Policies. 2015. 193 pp.

151 Selva Bahar Baziki. Firms, International Competition, and the Labor Market. 2015. 183 pp.

152 Fredrik Sävje. What would have happened? Four essays investigating causality. 2015. 229 pp.

153 Ina Blind. Essays on Urban Economics. 2015. 197 pp.

154 Jonas Poulsen. Essays on Development and Politics in Sub-Saharan Africa. 2015. 240 pp.

155 Lovisa Persson. Essays on Politics, Fiscal Institutions, and Public Finance. 2015. 137 pp.

156 Gabriella Chirico Willstedt. Demand, Competition and Redistribution in Swedish Dental Care. 2015. 119 pp.

157 Yuwei Zhao de Gosson de Varennes. Benefit Design, Retirement Decisions and Welfare Within and Across Generations in Defined Contribution Pension Schemes. 2016. 148 pp.

158 Johannes Hagen. Essays on Pensions, Retirement and Tax Evasion. 2016. 195 pp.

159 Rachatar Nilavongse. Housing, Banking and the Macro Economy. 2016. 156 pp.

160 Linna Martén. Essays on Politics, Law, and Economics. 2016. 150 pp.

161 Olof Rosenqvist. Essays on Determinants of Individual Performance and Labor Market Outcomes. 2016. 151 pp.

162 Linuz Aggeborn. Essays on Politics and Health Economics. 2016. 203 pp.

163 Glenn Mickelsson. DSGE Model Estimation and Labor Market Dynamics. 2016. 166 pp.

164 Sebastian Axbard. Crime, Corruption and Development. 2016. 150 pp.

165 Mattias Öhman. Essays on Cognitive Development and Medical Care. 2016. 181 pp.

166 Jon Frank. Essays on Corporate Finance and Asset Pricing. 2017. 160 pp.

167 Ylva Moberg. Gender, Incentives, and the Division of Labor. 2017. 220 pp.

168 Sebastian Escobar. Essays on inheritance, small businesses and energy consumption. 2017. 194 pp.

169 Evelina Björkegren. Family, Neighborhoods, and Health. 2017. 226 pp.

170 Jenny Jans. Causes and Consequences of Early-life Conditions. Alcohol, Pollution and Parental Leave Policies. 2017. 209 pp.

171 Josefine Andersson. Insurances against job loss and disability. Private and public interventions and their effects on job search and labor supply. 2017. 175 pp.

172 Jacob Lundberg. Essays on Income Taxation and Wealth Inequality. 2017. 173 pp.

173 Anna Norén. Caring, Sharing, and Childbearing. Essays on Labor Supply, Infant Health, and Family Policies. 2017. 206 pp.

174 Irina Andone. Exchange Rates, Exports, Inflation, and International Monetary Cooperation. 2018. 174 pp.

175 Henrik Andersson. Immigration and the Neighborhood. Essays on the Causes and Consequences of International Migration. 2018. 181 pp.

176 Aino-Maija Aalto. Incentives and Inequalities in Family and Working Life. 2018. 131 pp.

177 Gunnar Brandén. Understanding Intergenerational Mobility. Inequality, Student Aid and Nature-Nurture Interactions. 2018. 125 pp.

178 Mohammad H. Sepahvand. Essays on Risk Attitudes in Sub-Saharan Africa. 2019. 215 pp.

179 Mathias von Buxhoeveden. Partial and General Equilibrium Effects of Unemployment Insurance. Identification, Estimation and Inference. 2019. 89 pp.

180 Stefano Lombardi. Essays on Event History Analysis and the Effects of Social Programs on Individuals and Firms. 2019. 150 pp.

181 Arnaldur Stefansson. Essays in Public Finance and Behavioral Economics. 2019. 191 pp.

182 Cristina Bratu. Immigration: Policies, Mobility and Integration. 2019. 173 pp.

183 Tamás Vasi. Banks, Shocks and Monetary Policy. 2020. 148 pp.

184 Jonas Cederlöf. Job Loss: Consequences and Labor Market Policy. 2020. 213 pp.

185 Dmytro Stoyko. Expectations, Financial Markets and Monetary Policy. 2020. 153 pp.

186 Paula Roth. Essays on Inequality, Insolvency and Innovation. 2020. 191 pp.

187 Fredrik Hansson. Consequences of Poor Housing, Essays on Urban and Health Economics. 2020. 143 pp.

188 Maria Olsson. Essays on Macroeconomics: Wage Rigidity and Aggregate Fluctuations. 2020. 130 pp.

189 Dagmar Müller. Social Networks and the School-to-Work Transition. 2020. 146 pp.

190 Maria Sandström. Essays on Savings and Intangible Capital. 2020. 129 pp.

191. Anna Thoresson. Wages and Their Impact on Individuals, Households and Firms. 2020. 220 pp.

192. Jonas Klarin. Empirical Essays in Public and Political Economics. 2020. 129 pp.

193. André Reslow. Electoral Incentives and Information Content in Macroeconomic Forecasts. 2021. 184 pp.

194. Davide Cipullo. Political Careers, Government Stability, and Electoral Cycles. 2021. 308 pp.

195. Olle Hammar. The Mystery of Inequality: Essays on Culture, Development, and Distributions. 2021. 210 pp.

196. J. Lucas Tilley. Inputs and Incentives in Education. 2021. 184 pp.