ESSAYS IN LABOR AND EDUCATION ECONOMICS

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Alexander Lars Philip Willén
May 2018

ESSAYS IN LABOR AND EDUCATION ECONOMICS

Alexander Lars Philip Willén, Ph. D.

Cornell University 2018

ABSTRACT

This dissertation consists of three essays, each using advanced empirical methods to address important questions within the fields of labor and education economics.

In Chapter 1, I exploit a Swedish reform that eliminated the fixed national pay scale for teachers to present novel evidence on the labor market effects of wage decentralization. Identification of the causal effect of the reform is achieved by using differences in non-teacher wages across local labor markets prior to the reform as a measure of treatment intensity in a dose-response difference-in-difference framework. I find that decentralization induces large changes in teacher pay, and that these changes are entirely financed through a reallocation of existing education resources. The magnitude of the wage effect is negatively related to teacher age, such that the reform led to a disproportionate increase in entry wage and a flattening of the age-wage relationship. Contrary to the predictions of the Roy model, decentralization does not impact teacher composition or student outcomes. I show that a main reason for this relates to general equilibrium and wage spillover effects to substitute occupations.

In Chapter 2, which is joint work with Anders Böhlmark, we examine how ethnic residential segregation affects long-term outcomes of immigrants and natives. The key challenge with identifying neighborhood effects is that individuals sort across regions for reasons that are unobserved by the researcher but relevant as determinants of individual outcomes. Such nonrandom selection leads to invalid inference in correlational studies since individuals in neighborhoods with different population compositions are not comparable even after adjusting for differences in observable characteristics. To overcome this issue, we

borrow theoretical insight from the one-sided tipping point model used by Card, Mas and Rothstein (2008). This model predicts that residential segregation can arise due to social interactions in white preferences: once the minority share in a neighborhood passes a certain "tipping point," the neighborhood will be subject to white flight and avoidance, causing a discontinuity in white population growth. After having found evidence for the tipping phenomenon in Sweden, we use the tipping threshold as a source of exogenous variation in population composition to provide new evidence on the effect of neighborhood segregation on individual outcomes. We find negative effects on the educational attainment of native children. These effects are temporary and do not carry over to the labor market. We show that these transitory education effects are isolated to natives who leave tipped areas, suggesting that they may be driven by short-term disruptions caused by moving.

In Chapter 3, which is joint work with Michael Lovenheim, we analyze the effect of teacher collective bargaining laws on long-run labor market and educational attainment outcomes, exploiting the timing of passage of duty-to-bargain (DTB) laws across cohorts within states and across states over time. We find robust evidence that exposure to teacher DTB laws worsens the future labor market outcomes of men: in the first 10 years after passage of a DTB law, male earnings decline by $1,974 (or 3.64%) per year and hours worked decrease by 0.43 hours per week. The earnings estimates for men indicate that teacher collective bargaining reduces earnings by $198.1 billion in the US annually. We also find evidence of lower male employment rates. Exposure to DTB laws leads to reductions in the skill levels of the occupations into which male workers sort as well. Effects are largest among black and Hispanic men. Estimates among women are often confounded by secular trend variation, though we do find suggestive evidence of negative impacts among non-white women. Using data from the 1979 National Longitudinal Survey of Youth, we demonstrate that DTB laws lead to reductions in measured non-cognitive skills among young men.

BIOGRAPHICAL SKETCH

Alexander was born in 1989 to parents Bo Willén and Marianne Willén, née Andersson, in Ystad, Sweden. Alexander graduated from Red Cross Nordic United World College in Norway in 2008, after which he entered Durham University and received his B.A. in politics, philosophy and economics in 2011. He obtained a M.P.P from Georgetown University in 2013. In 2016 he received a M.S. in policy analysis and management from Cornell University, and in 2018 he was awarded his Ph.D. from the same university. Alexander is married to Adriana Condarco-Quesada.

ACKNOWLEDGMENTS

This document could not have been written without the help and support of many people, and I wish to express my warmest gratitude to all those persons who have influenced this work. I also wish to thank all those institutions that graciously provided me with financial support and data access during the work on this dissertation.

First and foremost, I owe a debt of thanks to my advisor Michael Lovenheim for his fundamental role in my doctoral work. His tireless support, patient guidance, and detailed feedback, have had a profound impact on my professional as well as personal development, and his passion for research has been a true source of inspiration. I am also indebted to the other members of my dissertation committee: Anders Böhlmark, Maria Fitzpatrick and Jordan Matsudaira. Their insightful comments and feedback were instrumental in shaping this document. I further thank Francine Blau and Lawrence Kahn. It has been wonderful to work with them for the past several years, and I have benefitted tremendously from their research insights and helpful advice.

In addition to help from my committee and the faculty and staff at Cornell University, I am grateful for the support I have received from the Institute for Evaluation of Labor Market and Education Policy in Uppsala and the Swedish Institute for Social Research at Stockholm University.

I thank my parents and siblings for their unconditional support and encouragement; my gratitude to them is beyond words. Lastly, I thank my wife Adriana, for making these the happiest times of my life, and without whose support this dissertation never could have been written.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

From a Fixed National Pay Scale to Individual Wage Bargaining: The Labor Market Effects of Wage Decentralization[a]

Abstract

Centralized wage-setting is a prevalent and controversial feature of the public sector labor market. This paper exploits a Swedish reform that eliminated the fixed national pay scale for teachers in 1996 to present novel evidence on the labor market effects of wage decentralization. Identification of the causal effect of the reform is achieved by using differences in non-teacher wages across local labor markets prior to the reform as a measure of treatment intensity in a dose-response difference-in-difference framework. I find that decentralization induces considerable changes in teacher pay, and that these changes are entirely financed through a reallocation of existing education resources. In terms of effect size, the results reveal a long-run teacher wage response elasticity of 0.2 with respect to the outside wage. The magnitude of this effect is negatively related to teacher age, such that the reform led to a disproportionate increase in entry wage and a flattening of the age-wage relationship. The analysis further shows that the policy increased wage dispersion among young and mid-career teachers. Contrary to the predictions of the Roy model, I do not find any economically significant effects on the composition of teachers or on student long-run outcomes. I show that a main reason for the lack of teacher composition and student outcome effects is due to general equilibrium and wage spillover effects to substitute occupations.

[a] I wish to thank Anders Böhlmark, Maria Fitzpatrick, Jordan Matsudaira and Michael Lovenheim for their advise and support. Francine Blau, Lawrence Kahn, Douglas Miller, Zhuan Pei, Amy Schwartz, Björn Öckert and seminar participants at Cornell University, Louisiana State University, Michigan State University, University of Oslo, the Norwegian School of Economics, the 2018 Association for Education Finance and Policy annual meeting, the 2017 Daniel Patrick Moynihan Summer Workshop in Education and Social Policy at Syracuse University, and the 2nd Stockholm-Uppsala Education Economics Workshop at Uppsala University, provided very helpful comments. I would further like to thank Pia Murphy at the Swedish Association of Local Authorities and Regions for providing detailed documentation on the centralized teacher wage negotiations that took place prior to the reform. I also thank Mimmi Råback and Jenny Davidsson at Lärarförbundet for information on the teacher wage scales prior to the reform. I gratefully acknowledge financial support from the Dr. Tech. Marcus Wallenberg Foundation [2015-040] and the Cornell University Graduate School.

1. Introduction

Despite a global trend toward greater wage decentralization over the past decades, centralized wage-setting remains a prevalent feature of the public sector labor market; the majority of OECD countries still pay postal workers, federal government employees, teachers and emergency services personnel according to wage scales that are held constant across often large and heterogeneous geographic areas.[1] A fundamental problem with this pay structure is that it cannot account for local labor market conditions. The regulated pay will therefore be relatively worse in areas with strong local labor markets, and existing research suggests that this may have adverse effects on the supply and quality of labor (Britton and Propper 2016; Propper and van Reener 2010). A more flexible pay system that allows wages to better reflect local labor market conditions may therefore improve labor supply and productivity.[2]

The central challenge to examining this question is a lack of plausibly exogenous variation in wage-setting regulation linked to detailed outcome data. I overcome this difficulty by exploiting a Swedish reform that abolished the fixed national pay scale for teachers in favor of individual wage bargaining in 1996. By combining several administrative registries, I follow individuals across time and space and trace the effect of the reform on teacher wage structure, teacher composition, and student long-run outcomes. To understand the mechanisms through which these effects operate, I evaluate the reform's effect on education spending, allocation of education resources, local tax rates, teacher-student ratios, as well as the general equilibrium effects of wage spillovers to other occupations.

---

[1] Between 1970 and 1990, the majority of OECD countries moved towards greater wage decentralization (OECD 2004). One exception to this was Norway, which witnessed a re-centralization of certain industries in the 80s (Kahn 1998).

[2] Centralized wage-setting is not exclusively associated with adverse labor market effects (e.g. Traxler 2003; Agell 1999; Harcourt 1997; Booth 1995; Calmfors 1993; Moene and Wallerstein 1993; Sapsford and Tzannatos 1993; Summers et al. 1993; Agell and Lommerud 1992; Layard et al. 1991; Soskice 1990; Calmfors and Driffil 1988; Blanchard and Summers 1986; Bruno and Sachs 1985; Crouch 1985). More broadly, wage centralization can reduce wage- and inflation pressure by preventing employers from closely related industries to engage in wage competition over labor. If wages are centralized across industries, it can promote expanding industries while forcing declining ones to leave the market. Increasing the size of the bargaining coalition through centralization may lead negative externalities that would exist under decentralization (such as passing on the cost of higher wages to consumers through higher product prices) to become internalized. Aidt and Tzannatos (2005) provides a thorough overview of the benefits and costs associated with centralized wage-setting, and Freeman and Gibbons (1995) explains in detail what led to the shift from centralized wage-setting to local bargaining in the 80s.

An enhanced understanding of the labor market effects associated with wage decentralization is of great general interest, and studying such effects in the teacher labor market is of particular importance. Even though existing studies find that teachers play a fundamental role in the determination of school quality (Slater et al. 2011; Aaronson et al. 2007; Clotfelter et al. 2007; Rivkin et al. 2005; Rockoff 2004; Goldhaber 2002; Darling-Hammond 2000; Hanushek et al. 1998), relative teacher pay has declined monotonically since 1940 (Hanushek and Rivkin 2007), and it is increasingly difficult to recruit and retain a sufficient stock of qualified teachers (OECD 2015; Corcoran et al. 2004; Hoxby and Leigh 2004).[3] Thus, teachers represent one of the most important inputs of the education production function, but current trends in compensation and supply make it unlikely that we will observe improvements in teacher quality except through policy interventions. Individual wage-setting has been proposed as one such intervention (Björklund et al. 2005), and if pay decentralization can reverse the above trends, it can have substantial implications for improving school quality.

Isolating the effect of the Swedish reform is complicated because it was implemented in the entire country at the same time. To overcome this difficulty, I use pre-reform variation in college-educated non-teacher pay across local labor markets (LLMs) as a measure of treatment intensity in a difference-in-difference framework. The intuition behind this approach is that the relative wage of teachers will be worse in areas with stronger local labor markets. The wage response to the policy should therefore be directly proportional to the outside wage.[4]

The identifying variation I use stems from cross-LLM differences in pre-reform college-educated non-teacher wages. The main assumption I invoke is that there are no secular trends, policies or shocks that affect outcomes differently depending on the area's pre-reform

[3] In Sweden, there was a substantial fall in relative teacher pay from 1950 to 1977. Since 1977, relative teacher pay has been held fairly constant, though there is some evidence of a slight decrease in the 80s and early 90s (Persson and Skult, 2014).

[4] The idea underlying this estimation strategy is similar to that in Card (1992), where he uses cross-state variation in the fraction of individuals that earn less than the 1990 Federal minimum wage as a measure of treatment intensity in a difference-in-difference model to examine the effect of legislated wage floors. More directly related studies that rely on a similar research design include Britton and Propper (2016) and Propper and van Reenen (2010). These studies are discussed in Section 3.

non-teacher wage. In addition to including a rich set of fixed effects and controls for other factors that may be correlated with both the pre-reform non-teacher wage and the outcomes I look at, I examine event-studies that explicitly test for the existence of pre-treatment trends in outcomes across cohorts. These results are inconsistent with plausible sources of bias from secular shocks or trends and support the causal interpretation of my estimates.

I find that wage decentralization induces considerable changes in teacher pay structure, and that these changes are entirely financed through a reallocation of existing education resources. In terms of effect size, the results reveal a long-run teacher wage response elasticity of 0.2 with respect to the outside wage. This result implies that a 10% increase in the pre-reform outside wage results in an increase of around 2% in teacher pay. Event studies show that this wage effect starts three years after the reform, increases gradually over time, and reaches a new equilibrium four years later. This delayed response is expected: local wage-setters cannot change the pay of workers overnight, especially given the limited pay guarantees that were in place during the first years after the policy change.[5] The magnitude of the wage effect is negatively related to teacher age, such that the reform led to a disproportionate increase in entry wage and a flattening of the age-wage relationship. My results further show that the policy led to modest increases in wage dispersion among young (20-34 years old) and mid-career (35-49 years old) teachers.

The Roy model (1951) asserts that workers choose professions based on the relative wage of the profession, the return to skill in the profession and the complementarity of skills across professions. Given the identified wage effects associated with the reform, the Roy model predicts the reform to raise the supply, and improve the quality, of teachers across regions that had higher pre-reform non-teacher wages relative to regions with lower pre-

[5] The limited guarantees ensured a $50 general increase for 1996 (an additional $39 was given to teachers that had reached the wage ceiling) and a minimum wage to teachers that had worked 1 and 5 years respectively beginning in 1997 (Söderström 2006). These guarantees are discussed in detail in Section 2.

reform non-teacher wages. In turn, such effects could improve student outcomes. My reduced-form estimates are at odds with these predictions: the policy did not have an economically significant effect on the supply and composition of teachers. I also do not find an effect on student long-run education and labor market outcomes. For example, I can rule out positive GPA effects larger than 0.07 national percentile ranks, and negative GPA effects larger than 0.7 national percentile ranks, from a 10% increase in the pre-reform non-teacher wage.

I show that a main reason for the lack of teacher supply and student outcome effects is due to wage spillovers to substitute occupations. Specifically, the identified increase in teacher wage makes it harder for firms outside the education sector to recruit and retain individuals with teaching qualifications. Further, it may induce individuals with non-teaching careers to consider teaching. This puts upward pressure on wages in these other occupations. I find strong support for a wage spillover effect to occupations closely related to teaching. This effect starts two years after the teacher wage effect and is half as large, reducing the relative teacher wage response to the reform by 50%. This spillover effect coupled with the compressed Swedish wage structure – which suggests that a response elasticity of 0.2 with respect to the outside wage only translates into small absolute teacher pay differences across municipalities post the reform – explains why the reform did not impact teacher composition and student outcomes. Specifically, I calculate that the long-run post-reform difference in absolute teacher wage across 68 percent of the municipalities, after accounting for the wage spillover effect, is less than 1.4 percent for females and 2 percent for males. It is unlikely that these small post-reform differences are sufficiently large to overcome existing search and matching frictions and mobility costs.

This paper makes several important contributions to the literature. First, I evaluate if decentralization induces municipalities to pay teacher wages that better align with local labor market conditions. This is an empirical question that no pre-existing study has been able to

examine. Second, I examine the reduced-form effect of wage decentralization on long-run teacher composition and student outcomes. The success of an education reform ultimately depends on how it impacts the long-run outcomes of students, and my ability to follow students across time and space represents a major contribution to the literature. Third, I examine how the reform affects education spending, allocation of education resources, local taxes and teacher-student ratios, providing direct evidence on the potential mechanisms through which the teacher composition and student outcome effects may operate. Finally, teachers represent the largest group of public sector employees in Sweden. As such, it is likely that changes to their wage structure impact wages in other occupations. I am the first to estimate these spillover effects.

In terms of policy implications, this paper shows that wage decentralization induces changes in wage structure financed entirely through a reallocation of existing education resources. These changes have no effect on teacher composition and student outcomes. A main reason for this is due to wage spillover effects to substitute occupations.[6] Thus, even if the government provides financial incentives to augment the local response to the decentralization reform, it still may not impact teacher supply and productivity because competing employers respond by also raising wages. The main takeaway from this paper is that wage decentralization induces a reallocation of education resources toward existing teachers. This has no effect on education quality, but leads to a broad cross-industry wage increase due to wage spillovers.

The rest of this paper proceeds as follows: Section 2 provides an overview of the Swedish education system, describes the reform, and offers a theoretical discussion on the likely effects of the reform. Section 3 reviews the literature, Section 4 introduces the data, and

[6] While the identified reallocation of existing education resources could also explain the lack of effects on teacher composition and student outcome, my results provide little evidence to suggest that spending shifted in ways that make teachers and students worse off. This is discussed in Section 7.

Section 5 presents the research design. Main results are shown in Section 6, Section 7 examines mechanisms, Section 8 presents results from diagnostic tests and robustness checks, and Section 9 concludes.

2. Institutional Background

2.1 The Swedish Education System

The Swedish education system consists of nine years of tuition-free comprehensive compulsory education starting at age 7, with the curriculum set by the central government. Following the completion of compulsory school, each child has the right to three years of tuition-free upper secondary school. In 2013, 98% of students that finished compulsory school continued to upper-secondary school (Skolverket 2014).

The majority of students attend public institutions; during my analysis period less than 1 percent of students attended private tuition-charging schools (Böhlmark and Lindahl 2015) and no more than 8 percent were enrolled at charter schools (Statistics Sweden 2006).[7] Children can choose to enroll at any school provided that space is available. However, individuals residing closest to the school are given priority at the grade school level, and proximity remains the main principle for allocating students to compulsory schools (Böhlmark and Lindahl 2015).

Since 1990, municipalities hold full financial responsibility for primary and secondary education, though substantial cross-municipality cooperation exists at the upper-secondary level.[8] To teach, an individual must hold an academic teaching degree. Traditionally, such a degree was obtained through a common university-level examination for all teachers. In 2011, this common exam was replaced with four specialized programs depending on which grades and subjects the teacher desires to teach. A school is not allowed to hire an unqualified teacher

[7] While there are no restrictions on the management structure of charter schools, they are not allowed to charge fees, must follow the national curriculum and are subject to the same rules for selecting students.
[8] Primary and secondary education in Sweden is funded primarily through a local municipality tax (70 percent) as well as through earmarked (5 percent) and non-earmarked (15 percent) grants from the national government.

if there is a certified teacher available for the position. Teacher pay negotiations remained centralized until the elimination of the pay scales in 1996. Decisions related to non-pecuniary benefits and work conditions remained at the national level throughout the analysis period.

2.2 Decentralization of Teacher Wages

A dominant feature of the Swedish labor market in the post-war era was its "solidarity wage policy" – equal pay for equal work (Edin and Holmlund 1995). This egalitarian ambition was pursued through wage bargaining via direct negotiations between peak associations (associations of industries or groups with shared interests or goals) of workers and employers (Iversen 1996). For employers, centralized bargaining ensured overall wage restraint and reduced competition for certain workers (Karlson and Lindberg 2011). For employees, it provided a more egalitarian wage distribution and guaranteed stable wage increases over time (Karlson and Lindberg 2008). This feature of the labor market induced substantial pay compression both across and within occupations: a 30% pay increase would move a worker from the bottom decile to the top decile of the wage distribution (Hibbs and Locking 1995). For an industrial worker in the US, the equivalent move would require a 400% pay increase (Hibbs and Locking 1995).

Concurrently with other OECD countries, Sweden moved away from centralized wage-setting in the late 70s (Granqvist and Regnér 2008). The transition toward individualized wage-setting began in the private sector as employers started questioning the centralized system: it gave little room for local and individual wage variation and made it difficult to recruit and retain high-quality workers (Granqvist and Regnér 2008). This idea soon gained support within the public sector as well, and Sweden is currently considered to have one of the most decentralized public sector wage-setting system in Europe (Ibsen et al. 2011).

The teacher pay decentralization reform was implemented in 1996, which was much later than the introduction of similar reforms in other public sector occupations in Sweden.[9] Prior to the reform, wages among primary and secondary teachers were determined through national pay scales based on type of teaching: (1) primary level, (2) lower-secondary level, (3) music/art/sport, (4) general subjects at the upper-secondary level, (5) vocational subjects at the upper-secondary level, and (6) lectureship. These scales determined entry wage, increases in wages with experience, and the maximum wage that could be received.[10] Figure 1 depicts the steps-and-lanes schedule for compulsory school teachers in 1990. On average, teachers enjoyed wage increases every 18 months for 15 years and then yearly for 5-8 years. After 20-23 years, a common wage ceiling was reached. Through negotiations between the national government and the central teacher unions (*Lärarnas Riksförbund* and *Lärarförbundet*), these scales were subject to upward shifts every 6 to 12 months (Söderström 2006).

There were two exceptions to the wage scales: teachers could receive small premiums for non-teaching duties (e.g. being responsible for gym equipment) and teachers in subjects that suffered from teacher shortages could receive small bonuses (Söderström 2006). Deviations to the pay scales based on these exceptions were very uncommon at the primary and lower secondary level – the focus of this analysis – and were generally considered restricted to teachers at the high school level. However, even at this level deviations were uncommon, and the single salary schedule should be considered deterministic prior to the reform (Söderström 2006).[11]

The decentralization reform was passed in the beginning of 1996. It was the result of careful negotiations between the two national teacher unions and the employer organization of teachers. Even though the employer organization was the main proponent of the reform, both

[9] For example, both nurses and doctors were introduced to individual wage-setting in 1989 (Calmfors and Richardson 2004).
[10] For non-certified teachers, the monthly wage was traditionally $200 less than the wage dictated by the pay scale.
[11] Though my analysis focuses on the primary and lower secondary level (i.e. compulsory school), the results for high school teachers are similar but noisier. This is further discussed in Section 4.

unions were dissatisfied with the single salary schedule because they felt it led to wage levels that were too low in general (Söderström 2006). The primary goal of the reform was to encourage wage differentiation based on performance and effort in order to recruit and retain high-quality teachers. Following the reform, teacher wages were determined through annual negotiations between individual teachers and their employers. However, the transition to individual wage-setting was gradual, and the new labor contract signed by the unions contained certain limited wage guarantees: a $50 general increase for 1996 (an additional $39 was given to teachers that had reached the ceiling) and a minimum wage to teachers that had worked 1 and 5 years respectively beginning in 1997 (Söderström 2006). This contract should therefore be seen as a transition contract.[12] The teacher labor contract was replaced again in 2000, at which point the transition was complete and the limited guarantees removed.

2.3 Theoretical Predictions

Following the existing literature on the labor market effects of centralized wage-setting (Britton and Propper 2016; Cardullo 2015; Propper and van Reenen 2010; Cappelli and Chauvin 1991), consider a dual-region model with *region = {L, H}*. Assume that the only difference between these regions concerns labor productivity, which is held constant across industries. Let L represent the low-productivity region and let H depict the high-productivity region. The unregulated non-teacher wage will be lower in L due to the lower productivity of the region.

At any given teacher wage, the difference in non-teacher pay across regions will cause teacher supply to be higher in L than in H. To equalize local supply and demand, the teacher wage needs to differ across the regions. This is not possible under centralized wage-setting,

[12] In addition to the limited guarantees, the transition contract allowed for two types of bargaining – between individual teachers and principals, and between local (municipality) labor unions and principals – and the pace of adoption of individual wage-setting is likely to differ depending on which of these two alternatives that local wage-setters opted for. Further, it may have taken time for both teachers and principals to develop bargaining skills and become comfortable with negotiating over wages. For a more detailed discussion on the particularities of the reform process and the actors involved, see Söderström (2006).

and the regulated pay creates a wedge between it and the local equilibrium wage that would prevail in the absence of the pay control. Figure 2 offers a visual depiction of this scenario.[13]

The resultant local labor market disequilibria may have negative effects on labor supply in H. First, the teaching profession will be more appealing in L due to the higher relative wage, and this may lead more, and more productive, workers to sort into teaching in L. Second, teachers in H may attempt to migrate to L in order to benefit from the higher relative wage. Third, the unemployment rate of teachers may be higher in L since the supply of teachers is higher in that region. Finally, worker productivity may be higher in L because a higher relative wage induces greater worker effort (Shapiro and Stiglitz 1984) and improves worker morale (Akerlof 1982).[14]

The above discussion suggests that local pay-setters may respond to the decentralization reform by starting to pay wages that better align with local competitive equalibrium pay. Such wage changes will likely lead to improvements in the supply and composition of teachers across regions that had higher pre-reform non-teacher wages relative to regions with lower pre-reform non-teacher wages. In turn, such effects could positively impact student outcomes.

A limitation of the theoretical framework discussed above is that it only considers how pay regulation affects wage levels. An equally important aspect of centralized wage-setting is that it restricts the return to skill in the profession, as everyone with the same level of education and experience are paid the same. Under wage decentralization, observed as well as unobserved individual characteristics that are more representative of individual productivity are more likely to enter the wage equation. This will raise the return to skill in the profession and likely augment the beneficial labor market effects of wage decentralization discussed

[13] Figure 2 assumes that the centralized wage falls right between the competitive local equilibrium wages in H and L. This is not a necessary assumption for the validity of the identification strategy used in this paper.

[14] It is worth noting that the centralized wage may not only have undesired consequences in H-type districts. In L-type districts the centralized wage acts as a price floor, and both teacher supply as well as teacher quality may exceed the optimal amount in these regions. The effect of the reform in these districts should be a relatively slower average growth rate of teacher wages over time (as sticky wages coupled with institutional features make it unlikely that these districts can cut teacher wages) and thus a relative reduction in teacher supply and quality.

above (Dahl et al. 2013). This assertion follows directly from the Roy model, which predicts occupational choice to be a function of not only the relative wage of the profession, but also of the return to skill in the profession and the complementarity of skills across professions.

There are three caveats accompanying the above predictions. First, local governments face budget constraints. Their ability to eliminate the pre-reform pay gap between local labor market and teacher wages will therefore depend on their capability to raise funds and reallocate resources across inputs and services. Thus, the magnitude of the reform effect on pay structure, which directly affects the labor supply response to the reform, will be a function of the ability of municipalities to reallocate resources toward teacher wages.

Second, even if local governments can obtain the necessary funds, the way in which they obtain these funds will directly impact the supply response to the reform. In theory, there are four ways through which municipalities can finance an increase in teacher pay: (1) lowering the teacher-student ratio, (2) reallocating resources across educational inputs, (3) raising the local income tax, or (4) redirecting resources from other parts of the public sector. If municipalities rely on (1), (2) or (3), the predicted effects of wage decentralization on labor supply and student educational attainment could be muted.[15]

Finally, an increase in teacher wage will make it harder for firms outside the education sector to recruit and retain individuals with teaching qualifications. Further, it may induce individuals with non-teaching careers to consider teaching. This will put upward pressure on wages in these other industries. The Swedish registry data coupled with the teacher wage decentralization reform provides me with the unique opportunity to examine such general equilibrium effects. This is an important contribution since no other study has been able to explore this question. Given the size of the teacher work force, the potential wage spillover effects are likely non-negligible. The existence of such spillover effects would reduce the

[15] If they rely on (4), the decentralization reform may have negative effects on other parts of the public sector. However, the results presented in Section 7 provide no evidence of municipalities resorting to this option.

impact of the reform on teacher supply and productivity, and this part of the analysis is therefore imperative for understanding the net effect of wage decentralization.

Taken together, the above discussion predicts that the reform will induce local wage-setters to pay wages that better align with local labor market conditions. All else constant, the Roy model predicts that this will have positive effects on the supply and quality of teachers in districts with high pre-reform non-teacher wages relative to districts with low pre-reform non-teacher wages. In turn, this could positively impact student outcomes. However, any wage response to the reform represents an increase in the cost of teachers. The above discussion makes clear that there also may be countervailing spending and resource allocation effects that obscure the labor market predictions of the Roy model. Potential general equilibrium and wage spillover effects add to the difficulty of anticipating the likely consequences of the reform on teacher composition and student outcomes. The net impact of the reform on teacher supply, productivity and student outcomes is thus ambiguous, underscoring the importance of this empirical analysis.

3. Prior Literature

The central challenge facing the existing wage decentralization literature is a lack of exogenous variation in pay-setting regulation. The majority of research on this topic is therefore cross-sectional, leveraging variation in individual wage-setting power across and within industries at a given point in time (Daouli et al. 2013; Fitzenberger et al. 2013; Granqvist and Regnér 2008; Fitzenberger et al. 2008; Plasman et al. 2007; Dell'Aringa and Pagani 2007; Card and de la Rica 2006; Gerlach and Stephan 2005; Cardoso and Portugal 2005; Rycx 2003; Dell'Aringa and Lucifora 1994a). This literature suggests that decentralization is associated with a large wage premium. Card and de la Rica (2004), for

example, estimate the premium to be about 10 percent. However, the literature fails to reach a consensus on how wage dispersion is affected.[16]

A major limitation with the cross-sectional studies is the likely existence of unobserved heterogeneity in worker productivity. If pay decentralization affects wage structure, then there may be sorting across occupations with and without local pay bargaining that bias the results. For example, if local wage bargaining generates increased wage dispersion and raises the pay level, then high ability workers may self-select into occupations with local wage bargaining. This would lead to an upward bias of the effect of wage decentralization on pay level. More recent studies have been able to account for this heterogeneity by using longitudinal employer-employee matched data and find that the pay premium associated with individual wage bargaining becomes smaller (2-4 percent) – but remains statistically significant – once this heterogeneity is accounted for (Andréasson 2014; Dahl et al. 2013; Gurtzen 2007). It should be noted that these papers focus on the wage effect of private sector pay decentralization brought about by a general shift in employer attitude. It is not clear that the wage effect of a government-mandated decentralization reform in the public sector should be the same.

The three papers most closely related to the current analysis are Britton and Propper (2016), Biasi (2017), and Söderström (2006). Britton and Propper (2016) study the effect of centralized wage-setting on education quality. To examine this question, the authors note that teacher pay is held largely constant across England, but that relative teacher pay varies due to regional differences in private sector wages. The authors exploit the wedge between the regulated teacher wage and the non-regulated outside wage across regions to analyze the effect of centralized wage-setting through a difference-in-difference design.[17] They find that a 10%

[16] Dell'Aringa and Lucifora (1994b) find that wage dispersion goes down, Card and de la Rica (2006) and Rycx (2003) find that wage dispersion goes up, and Plasman et al. (2007) and Dell'Aringa and Pagani (2007) find mixed results.

[17] Propper and van Reenen (2010) rely on the same identification strategy to examine the effect of centralized wage-setting on the quality of nurses in the UK. The results suggest that the number of hospital deaths within thirty days of emergency admission for acute myocardial infarction is about 6% higher in regions where the outside wage is 10% higher.

increase in the wage gap between local labor market and teacher wages reduce school performance in key exams at the compulsory level by 2%.

While it is important to identify the labor market effects of centralized wage-setting, it is equally important to understand how the elimination of such a schedule impacts the composition and supply of labor. This cannot be inferred from Britton and Propper (2016), because decentralization may affect not only relative pay, but also pay dispersion, spending, resource allocation, teacher-student ratios and local taxes. Further, it may have wage spillover effects to other occupations. Thus, it is unlikely that the effect of decentralization is symmetric to the effect of centralized wage-setting. My contribution to this literature is to examine if the adverse effects of centralized wage-setting can be eliminated through decentralization.

Biasi (2017) examines the labor market effects of Wisconsin's Budget Repair Bill of 2011. This bill aimed to resolve the 3.6 billion dollar budget deficit that the state was facing by changing several public sector employee regulations within the areas of collective bargaining, compensation, health insurance and sick leave. One of the implications of this bill was that districts could now choose to negotiate individual wages with teachers if they so desired, thus allowing districts to move away from a single salary schedule to individual wage bargaining. Using a difference-in-difference model that compares teacher in districts that chose individual bargaining to teachers in districts that decided against individual bargaining, Biasi finds that pay dispersion increased in districts that implemented individual bargaining and that this increase is positively correlated with teacher value-added.

An important difference between Biasi (2017) and the current study is that I investigate the effects associated with a government-mandated national wage decentralization reform, while the analysis in Biasi is based on voluntary adoption of individual wage-setting by districts made possible through a Bill that affected much more than only the salary schedule of teachers. While both settings are interesting, my analysis is more likely to isolate the effect of

wage decentralization. Further, the Swedish registry data permit a more comprehensive analysis of the labor market effects of wage decentralization. For example, these data provide information on which occupations teachers come from and to which occupations they leave, allowing me to examine both sorting and wage spillover effects. In addition, I can examine how the reform affects the long-run outcomes of students. The relative success of an education reform ultimately depends on how it impacts these outcomes, and my ability to follow students across time represents an important contribution to the literature.

The third paper that is closely related to the current analysis is Söderström (2006), which uses a difference-in-difference framework to compare the wages of Swedish teachers to that of other public sector employees before and after the teacher wage decentralization reform. The results suggest that the reform led to an increase in entry-wage, that the age-earnings profile became flatter and that earnings dispersion increased for old teachers. A limitation with Söderström (2006) is that the author only examines the earnings effect of the reform. I contribute to this literature by examining how the reform affects the composition of teachers and the long-run outcomes of students, as well as by investigating the mechanisms through which these effects operate. It is also important to note that I find substantial wage spillover effects associated with the reform, which suggests that the results in Söderström (2006) may be attenuated since his control group also was affected by the reform.

In addition to contributing to our understanding of the effects of wage decentralization, my paper is also related to several strands of the economics of education literature. One of these strands focuses on the effect of wages on teacher supply. The results from these studies are in line with conventional labor theory (Guarino et al. 2006): teacher wage is positively associated with retention and inversely related to attrition (Clotfelter et al. 2008; Imazeki 2005; Podgursky et al. 2004; Hanushek et al. 2004; Stockhard and Lehman 2004; Lankford et al. 2002; Kirby et al. 1999; Weiss 1999; Brewer 1996; Gritz and Theobald 1996), improving

teacher pay raises a district's ability to recruit high quality teachers (Leigh 2012; Figlio 2002), and higher wages lead to an increase in teacher supply (Falch 2010). If pay decentralization leads to higher wages, these studies suggest that the reform will improve teacher quality. However, since the reform also may affect other components of the education production function (e.g. through a reallocation of education resources) and broader labor market, this prediction may not come true.

Another body of research investigates the effect of teacher wages on student outcomes. The earlier studies within this field fail to identify statistically significant effects, suggesting that increases in teacher pay do not improve student outcomes (Hanushek 2003; Hanushek 1997; Grogger 1996; Betts 1995; Altonji 1988). In an influential paper, Loeb and Page (2000) offer another explanation: the earlier papers are unable to isolate the effect of interest because they fail to account for alternative labor market opportunities and non-pecuniary school characteristics. Adjusting the empirical method to account for local labor market factors, Loeb and Page (2000) finds that a 10% increase in teacher pay reduces high school dropout rates by 4% and increases college enrollment by 1.6%. More recent papers have produced similar results (e.g. Hendricks 2014; Dolton and Marcenaro-Gutierrez 2011). If decentralization leads to an increase in teacher pay, then this strand of research would predict the reform to also improve student educational attainment. However, this assumes that the reform does not affect other components of the education sector and broader labor market in ways that offset the effects of higher wages.[18]

To summarize, existing research largely agrees on the effect of individual wage-setting on the level of pay, but fails to reach a consensus on how it affects wage dispersion. There is

[18] There is also a literature on how teacher supply affects student outcomes. The results from these studies are in line with theoretical predictions: teacher turnover has a disruptive impact on student performance, experienced teachers have a large positive effect on student achievement, and high quality teachers have a positive impact on student outcomes (Chetty et al. 2014; Ronfeldt et al. 2011; Staiger and Rockoff 2010; Clotfelter et al. 2007; Boyd et al. 2006; Rivkin et al. 2005; Rockoff 2004). If the reform has a positive impact on teacher supply, these studies would therefore anticipate positive effects on student outcomes as well. However, it is unlikely that the reform only affects teacher supply, making it inappropriate to hypothesize the effects of the reform based only on this strand of research.

no research on the effects of wage decentralization on spending and resource allocation, or on whether decentralization in one occupation has spillover effects to other occupations. Thus, even though the economics of education literature produces relatively clear predictions on the supply response to wage level changes, this is not sufficient for identifying the labor market effects of wage decentralization. Prior to this study, we do not yet have a full understanding of the dynamics and general equilibrium effects of teacher wage decentralization.

4. Data

This paper uses administrative data from 1991 to 2006 drawn from several registries of the Institute for Evaluation of Labor Market and Education Policy database, originally collected by Statistics Sweden. The first registry is the *Teacher Registry*, which contains annual labor market information on all teachers in Sweden (workplace, type of contract, what they teach and if they are on leave).[19] I complement these data with the *Wage Registry for Public Sector Employees*, which contains information on wage and occupation, as well as work hours, for every public sector employee in Sweden. I further supplement these data with the *Teacher Education Registry*, which provides information on the education that each teacher has.

Although the reform affected all teachers in grades 1 through 12, I focus on teachers in grades 1 through 9. I impose this restriction due to certain institutional features. First, several municipalities are members of cross-municipality associations that share the responsibility of providing upper-secondary education to its residents. These associations vary in size, across time and with respect to the particularities of the agreements. This means that wage decisions are not made at the municipality level, but at the association level through negotiations between the involved municipalities, and the municipalities change over time.[20] Second, as discussed Section 2, the two exceptions to the wage scales were more common at the upper-

[19] The data do not include individuals over the age of 65. However, only 2.5% of teachers were over the age of 65 in 2014, and this was the highest percentage over the period 1990-2014 (Hansson 2015).
[20] It is also the case that municipalities with very few students do not have high schools and pay for their students to attend high schools in larger municipalities nearby in which they have no control over wage decisions.

secondary level. However, it should be noted that the results for high school teachers are similar to those for elementary school teachers, though the standard errors are larger.[21]

Excluding teachers in leadership roles, there is an average of 85,000 public elementary school teachers per year between 1991 and 2006, giving me more than 1.3 million teacher-year observations for the purpose of this analysis. The first two columns of Table 1 provide summary statistics of key socioeconomic and demographic variables for teachers and non-teachers respectively. Table 1 shows that teaching in public elementary schools is female-dominated - only 26 percent of teachers are male. This is consistent with the gender composition of teachers across the industrialized world. The table also shows that the teacher workforce is relatively old. The mean age is 45.4, compared to 41.3 in other sectors. Another feature of the teacher labor market is that teachers are highly educated. The average years of schooling among elementary school teachers is 14.9. This is much higher than that of non-teachers (11.6). Finally, almost 64 percent of teachers are married, and 59 percent have at least one child under the age of 18. These numbers are noticeably higher than the averages in other sectors, which suggests that geographic labor mobility may be lower among teachers than among workers in general.

Table 1 also provides information on these characteristics for specific time-periods: the pre-reform period (91-95), the direct post-reform period (96-00) and the more distant post-reform period (01-06). Some trends are worth noting. First, there is a reduction in the fraction of permanent teachers, fewer teachers hold teaching certificates and they work fewer hours. Second, the family composition of teachers appears to change, with fewer teachers getting married and having children. In terms of wage, there is a large increase in the pay level.

In Section 2, I note that centralized wage-setting prevents local wage-setters from paying the local competitive equilibrium wage and that this may negatively impact teacher

[21] These results are available from the author upon request.

supply in municipalities with high non-teacher pay. To examine if my data is consistent with this hypothesis, Table 2 compares the composition of teachers in municipalities in the bottom decile of the college-educated non-teacher employment income distribution with that of teachers in municipalities in the top decile in the year prior to the reform.

Table 2 suggests that the composition of teachers in municipalities in the top decile of the college-educated non-teacher employment income distribution is significantly different from that in municipalities in the bottom decile. Specifically, they are 3.2% less likely to remain in the profession, have a 1.8% higher average age, are 8.1% less likely to hold a teacher certificate, have 1.7% fewer years of schooling, have a 28.5% greater chance of being hired on a temporary basis and are 6.1% more likely to be female. Although simple differences in means cannot be used for causal inference, the results are consistent with both theory and prior literature.

The rich teacher data permit a very detailed analysis of how the decentralization reform affected teacher composition. The outcome characteristics I look at include age, years of schooling, master's degree, immigrant status, gender, being on leave, the probability of switching from private to public school, the probability of being on a temporary contract, hours worked (measured as a percent of a 40 hour work week), certification status, the probability to move to a different municipality, the probability to move to a municipality that had a higher pre-reform non-teacher wage, retention, and recruitment. Panel A of Appendix Table A-1 shows summary statistics of these variables.[22] Despite the high education level of teachers, the monthly mean wage of public elementary school teachers is low ($2,806), and Swedish teachers are in the left-tail of the OECD teacher pay distribution (OECD 2016).[23]

[22] The wages are expressed in real 2005 values.
[23] Raw wage comparisons are misleading as there are important cross-country differences in working conditions. For example, Swedish teachers had less than two-thirds as many teaching hours as teachers in Switzerland, and were responsible for half as many children as teachers in Ireland, where wages were much higher during the analysis period (OECD 1996).

I also use the *Longitudinal Database for Education, Income and Labor Market Participation* (*LOUISE*) registry. These data contain annual socioeconomic and demographic information on all residents between the ages of 16 and 65. The data include education, labor market, income and welfare program information. I use these data for three purposes. First, to obtain municipality-specific covariates that, if omitted from the model, could confound the estimates. Second, to obtain the long-run labor market outcomes of students that were exposed to the reform while in compulsory school. Finally, to obtain the treatment variable – the employment income of college-educated non-teachers across LLMs. This variable differs from the wage variable for teachers as it includes compensation for sick leave and for commuting to and from work. I rely on this measure because the *Wage Registry for Private Sector Employees* (the private sector equivalent of the *Wage Registry for Public Sector Employees*) is a survey based on a random sample covering less than 50% of private employees, suffers from nonresponses that cannot be explained and is subject to a stratification method that leads to a disproportionate loss of observations in small municipalities.[24]

I use Statistic Sweden's classification of LLMs. Specifically, a municipality is considered the center of a LLM if less than 20% of its working-aged population commute to a job outside the municipality and no more than 7.5% of the working-aged population commute to one specific outside municipality for work. Municipalities that do not meet these restrictions are allocated to the LLM to which the majority of its working-aged population commutes. There is at least one municipality in each local labor market, and no municipality is associated with more than one local labor market. In 1995, there were 106 LLMs and 288 municipalities in Sweden.

[24] This choice is unlikely to affect the results as the correlation between employment income and wage is above 0.9 among college-educated non-teachers that are present in both registries. To bias the results, it would further have to be the case that there are systematic differences between employment income and wage across individuals in municipalities that have different college-educated non-teacher wages in the year prior to the reform, and this is unlikely. It is worth noting that the estimates presented in this paper are robust to using teacher employment income rather than wage. These results are available on request.

The gender-specific mean employment incomes of college-educated non-teacher males and females in 1995 were $39,954 and $25,205, respectively.[25] The average difference between the college-educated outside employment income and the teacher wage in a municipality was $828 among males and -$178 among females. These values represent 2.1 and -0.7 percent of the employment income of college-educated non-teacher males and females prior to the reform. In no municipality was the male teacher wage higher than the associated LLM non-teacher college-educated employment income. Female teachers enjoyed higher wages than their non-teaching counterparts in almost 90% of the municipalities. The large cross-gender difference in the wedge between the outside non-regulated employment income and the teacher wage is driven entirely by differences in college-educated non-teacher employment incomes; the average difference between the monthly male and female teacher wage across municipalities was only $65 in the year prior to the reform.

Figure 3 provides a visual depiction of the cross-LLM variation in pre-reform non-teacher employment income with respect to males and females respectively. In the Figure, LLMs have been color-coded based on which decile of the gender-specific non-teacher employment income distribution they belong to, with LLMs in yellow belonging to the bottom decile and LLMs in brown belonging to the top decile. Black solid lines indicate 1995 LLM borders, while gray solid lines indicate municipality boarders. All LLM borders are also municipality borders.

Looking across Figure 3, there is substantial geographic variation in the treatment variable. There are also large gender differences in the geographic variation of pre-reform non-teacher employment income. The correlation between the treatment variable for males and females is only 0.524. This is encouraging as it makes it unlikely that the reform coincides

[25] While the female mean is very similar to that in the US ($24,555), the male mean is noticeably smaller ($49,928). Information on US employment income has been taken from Census (1996). These numbers have been deflated to represent real 2005 values.

with other shocks or policies isolated to municipalities with particular pre-reform non-teacher employment incomes that might influence post-reform teacher wages.

To examine the impact of the reform on short-term student educational outcomes, I use the *Grade 9 Registry*, which provides information on the academic performance of individuals in 9th grade including GPA and individual grades in the core subjects of math, Swedish and English.[26] I track these students through high school (via the *High School Registry*) and into the labor market ten years after graduating from 9th grade (via *LOUISE*) to examine long-term education and labor market effects of the reform. The education outcomes I look at in this part of the analysis are high school GPA, whether the student attended a university-preparatory high school track, graduated from a high school science program, and was ever enrolled at university. The labor market outcomes I study are the probability of being in the earnings sample, the probability of being a social security recipient, employment income, government-funded benefits (compensation from 32 social security programs, including educational grants, unemployment benefits, early-retirement supplemental compensation, compensation for start-ups and compensation for voluntary military service), and social insurance benefits (income from a set of social security programs for which participation is conditioned on employment). Summary statistics are provided in Panel B of Appendix Table A-1.

In addition to the registry data described above, I rely on three public-use data sets released by the Swedish National Agency for Education (SNAE) and Statistics Sweden (SCB). First, SNAE releases municipality-specific information on education spending stratified by input (teaching, food, facilities, health, supplies and uncategorized items). I use this information to examine if the reform affected education spending and resource allocation.[27]

[26] During my analysis period, Statistics Sweden did not collect statistics on grades until the students reached 9th grade.

[27] The inputs for which spending is reported vary by year, and I use only the categories that are consistently measured throughout the analysis period. Per-student spending on uncategorized items has been constructed to equal the difference between total per-student spending and per-student spending on teaching, food, health, supplies and facilities. This measure accounts for any spending that did not fall into any of the categories that municipalities were asked to report spending on, and any spending that falls into categories that municipalities were asked to report spending on for only a subset of the years under examination. These

Appendix Table A-2 provides summary statistics on per student spending by educational input. Second, SNAE also releases municipality-specific information on the number of students and the fraction of students enrolled at charter schools. I use this information to control for potential variation in public school cohort size that could confound my estimates. This information further enables me to investigate if the teacher-student ratio was affected by the decentralization reform. Third, SCB publishes information on local tax rates. I use this information to examine if the potential wage effect was funded, at least in part, through an increase in the municipality income tax. The SNAE data are only available beginning in 1992.

5. Empirical Methodology

I exploit cross-LLM variation in pre-reform college-educated non-teacher employment income as a measure of treatment intensity in a difference-in-difference framework. Specifically, I estimate models of the following form:

$$lnTW_{gmlt} = \alpha + \beta_1\left(lnNTW_{gl,1995} * Post_t\right) + \gamma X_{gmlt} + \delta_{gm} + \theta_{gt} + \varepsilon_{gmlt}, \quad (1)$$

where $lnTW_{gmlt}$ is log monthly wage of public elementary school teachers of gender g in municipality m and LLM l at time t. $lnNTW_{gl,1995}$ is log monthly employment income of college-educated non-teachers of gender g in LLM l in the year prior to the reform, $Post_t$ is an indicator variable equal to 1 if the observation is from the post-reform period and X_{gmlt} are municipality covariates that include both socioeconomic and demographic characteristics.[28]

The unit of observation is a municipality-gender-year, providing me with a balanced panel of 576 observations per year. Aggregation to this level is sensible because municipalities

categories include, but are not restricted to, cost of school library, career services, administration, student transportation, home language instruction and Swedish for immigrants.

[28] Provided that there were no systematic differences in teacher wages prior to the reform across municipalities, using $lnNTW_{gl,1995}$ as the treatment variable should be very similar to using $(lnNTW_{gl,1995} - lnTW_{gl,1995})$. In results not shown I have reestimated equation (1) using this alternative treatment measure, and my baseline results are insensitive to this modified treatment measure. These results are available upon request.

are responsible for education at the elementary school level, and males and females face disparate labor market opportunities. The identifying variation stems from gender-specific differences in college-educated non-teacher employment income across LLMs in the year prior to the reform.

The treatment variable of interest is $lnNTW_{gl,1995} * Post_t$. β_1 is the response elasticity of teacher wage with respect to the pre-reform college-educated non-teacher employment income. This coefficient captures the effect of the reform that can be explained by variation in pre-reform college-educated non-teacher employment income across LLMs. Equation (1) also includes a set of gender-by-municipality (δ_{gm}) and gender-by-time (θ_{gt}) fixed effects. The former controls for variation in teacher pay that is common to all teachers of a specific gender within a municipality over time. The latter controls for variation in teacher pay that is common to all teachers of a particular gender across municipalities in a given year.[29]

The main assumption underlying the identification of parameter β_1 is similar to that in all difference-in-difference analyses: the reform must be uncorrelated with prior trends in teacher wages across municipalities with different non-teacher employment incomes in 1995. I estimate the following event-study model to show that the data support this assumption:

$$lnTW_{gmlt} = \beta_0 + \sum_{t=1991}^{2006}[\pi_t(lnNTW_{gl,1995})] + \gamma X_{gmlt} + \delta_{gm} + \theta_{gt} + \varepsilon_{gmlt}, \qquad (2)$$

where π_t is the effect of college-educated non-teacher employment income in the pre-treatment year on teacher wage in year *t*. This relative time parameter should be flat and not statistically significantly different from zero in the pre-reform period.

A key benefit of equation (2) is that it allows me to relax the time-invariant treatment assumption underlying estimation of β_1 in equation (1). This is important as the wage

[29] It is worth noting that I am unable to disentangle nation-wide level effects of the reform from the year fixed effects: Should the reform lead to an overall level shift in teacher supply this will be subsumed by θ_{gt} and is therefore not something that can be picked up by β_1.

response to the reform may vary over time. Wage-setters cannot change the pay of workers overnight, especially given the limited guarantees that were in place during the transition period of 1996-2000. By non-parametrically tracing out the full adjustment path of the treatment effect via equation (2), I am able to examine the dynamic response to the reform.

In addition to the parallel trend assumption, the validity of my results require that the reform does not coincide with any shocks or policies isolated to municipalities with particular pre-reform non-teacher employment incomes that might influence post-reform teacher wages. Although it is difficult to directly investigate this assumption, I note that there is substantial cross-LLM variation in pre-reform college-educated non-teacher employment income. After controlling for fixed differences across municipalities and over time, and a rich set of time-varying municipality characteristics, it is unlikely that there are secular shocks in 1996 that are systematically correlated both with the employment income of non-teacher college-graduates in 1995 and the outcomes I examine.

6. Results

6.1 The Effect of Wage Decentralization on Wage Structure

The effect of wage decentralization on teacher pay is shown in Figure 4 (a), obtained from estimating equation (2) using all public elementary school teachers in Sweden. Each dot is an estimate of relative time parameter π_t for the given year. The bars extending from each point show the bounds of the 95% confidence interval. Both the treatment and dependent variable are measured in logarithmic form. π_t therefore represents the response elasticity of the teacher wage with respect to pre-reform non-teacher employment income.

Figure 4 (a) shows a clear wage effect associated with the decentralization reform, and several observations are worth highlighting.[30] First, wages are trending similarly across

[30] In terms of interpreting these results, it is important to note that the teacher wage did not decline in any of the municipalities. Rather, it increased differentially across the municipalities as a function of the pre-reform college educated non-teacher employment income.

municipalities in the pre-period as a function of the pre-reform non-teacher employment income. The figure is therefore inconsistent with the existence of pre-treatment trends that could bias the results. Second, it takes three years for the wage to react to the reform. This lag is expected: local wage-setters cannot change the pay of its workers overnight, especially given the limited guarantees that were in place during the transition period of 1996-2000. Third, the treatment effect grows over time until it levels out seven years after the reform. Interpreting this point as the stable long-run treatment effect, the figure suggests a long-run response elasticity of teacher wage with respect to pre-reform non-teacher employment income of about 0.2.[31] The reform thus induces local wage-setters to pay wages that better align with local competitive equilibrium pay. However, the effect size is likely not sufficiently large to equalize relative teacher pay across municipalities.[32] This incomplete response is suggestive of wage-setters being unable (due to budgetary constrains) or unwilling (due to competing priorities) to fully eliminate the difference in relative teacher pay across local markets. In terms of policy implications, if the goal of decentralization is to equalize relative pay across districts, it may be necessary for the government to provide financial incentives to local wage-setters to encourage a more substantial wage response.[33]

---

[31] Given the noticeable cross-gender difference in the treatment measure, it is possible that there are asymmetric treatment effects with respect to gender. This is examined in Appendix Figure A-1. Even though the effect size is slightly larger for women, this difference is not statistically significant.

[32] Assuming that teacher quality is constant across municipalities, equalization of relative teacher pay would require a response elasticity of 1. However, as shown in Table 2, this assumption likely does not hold. Yet, Table 2 also shows that the differences in teacher characteristics across the municipalities are relatively small, and it is therefore unlikely that a response elasticity of 0.2 equalizes teacher pay across municipalities conditional on teacher quality.

[33] There are a few institutional features, trends, and labor market reforms that occurred during my analysis period that could potentially bias the results shown in Table 4. First, *Skolvalsreformen* of 1992, which allowed for-profit charter schools to enter the education market. If the college-educated non-teacher employment income in 1995 is correlated with the growth of for-profit charter schools in the region, that may bias me towards finding positive wage effects. The reason is that the entry of for-profit schools leads to increased competition over labor and potentially higher wages. However, my results are robust to controlling for both the fraction of students in charter schools over time and its interaction with my treatment variable. Second, *Kunskapslyftet*, a government-mandated program that ran between 1997 and 2002, and focused on providing individuals with less than a high school degree complementary education. This may have increased the demand for teachers, and if this is correlated with the college-educated non-teacher employment income in 1995, it could bias my results. To investigate this potential confounder, I include as a control variable the number of people with less than a high school degree, and its interaction with the treatment variable between 1997 and 2002, and reestimate equation (1). My results are robust to the inclusion of these controls. Third, the Balkan war led to a substantial increase in immigration to Sweden between 1993 and 1996, and may have put increased pressure on the demand for teachers and their wages. To investigate this potential confounder, I include as a control variable the number of immigrants, and its interaction with the treatment variable between 1993 and 1996. My results are robust to adding these controls. Finally, there may be differential trends in the demand for teachers due to demographic shifts during the time period considered. To investigate this potential source of bias, I include as a control variable the total number of students in the municipality, and its

The effect shown in Figure 4 (a) could mask substantial treatment heterogeneity across the teacher wage distribution. In other panels of Figure 4, I therefore also show how the reform affected (b) the median wage, (c) the 10th percentile wage, (d) the 90th percentile wage, (e) the interquartile range, and (f) the standard deviation.[34] The dynamics of the median wage effect mirrors the impact of the reform on the mean teacher wage, but the magnitude of the effect is marginally larger. This is indicative of the reform causing slight overall tightening of the wage schedule and stands in contrast to one of the goals of the reform - to increase wage dispersion as a means to improve productivity through differential wages based on performance. Figures 4 (c) and (d) show that this wage compression is due to the reform having a greater wage effect at the left-tail of the distribution, with a point elasticity twice as large as that in the top decile.[35] Figure 4 (e) shows a statistically and economically significant negative effect of the reform on the interquartile range, providing direct evidence of a tightening of the wage distribution. Similar to the wage level effect, the effect on the interquartile range is time-varying and reaches a new long-run equilibrium seven years after the reform. However, the magnitude of this effect is small: a 10% increase in pre-reform non-teacher employment income leads to a reduction in the interquartile range of $16.18. The standard deviation appears unaffected. Results obtained from estimation of equation (1) for each of the above outcomes are shown in Table 3. Although the estimates depicted in Table 3 are consistent with the event studies in Figure 4, they are slightly attenuated due to the inability of equation (1) to account for time-varying treatment effects.

The effects shown in Table 3 - an overall reduction in wage dispersion driven by larger wage level effects at the left-tail of the wage distribution - are consistent with Söderström

---

interaction with the treatment variable, and reestimate equation (1). Adding these controls does not have a statistically significant effect on my point estimate. All these results are available from the author upon request.

[34] The wage dispersion measures have not been subject to log transformations because the pre-reform wage distribution is very compressed, in particular for old teachers that have reached the wage ceiling. Regressions stratified by age can therefore not be performed with log transformed dispersion measures as several of the municipalities have values of zero. For consistency, none of the dispersion results discussed in this paper therefore come from regressions with log transformed dispersion measures. However, all non age-specific wage dispersion estimates are robust to this adjustment.

(2006) who finds suggestive evidence of the reform causing a disproportionate increase in teacher entry wage and a flattening of the age-wage relationship. This result is also consistent with the idea that the pre-reform schedule compressed entry wage and provided an above-market return to experience among teachers. Before the reform, a 50-year-old teacher earned approximately 50 percent more than a 26-year-old teacher, while a 50-year-old non-teacher earned only about 20 percent more than a 26-year-old (Söderström 2006).[36] Thus, a flattening of the age-wage relationship is consistent with local wage-setters adjusting the teacher pay structure to better align with the non-regulated market rate of return to experience.

The pooled results in Table 3 may hide substantial treatment heterogeneity across different teacher cohorts, and I therefore reestimate the wage effects separately for young (20-34 years old), mid-career (35-49 years old) and old (50-64 years old) teachers. Results from estimation of equation (1) are shown in Table 4. Due to the relatively small size of some of the municipalities coupled with the facts that less than 27% of teachers are males and that the average teacher age is 45, there will not be a complete set of gender-municipality-year cells for each of these cohorts.[37] The pooled results will therefore not be equal to the average of the cohort-specific results.

The results in Table 4 show that the wage effect is negatively related to teacher age. This supports the idea that the reform led to a disproportionate increase in teacher entry wage and, as a consequence, a flattening of the age-wage relationship. Table 4 also shows that the overall wage compression effect identified in Table 3 hides substantial treatment heterogeneity across age: A 10% increase in pre-reform non-teacher employment income leads to a $9.3 increase in the standard deviation of the monthly wage among young teachers (2.9 percent

[36] More generally, existing literature on the return to experience suggests that there are no additional gains, or that the gains are substantially reduced, after 10 years of experience (e.g. Heckman et al. 2006). As shown in Figure 1, this was not true for Swedish teachers before the reform.

[37] 8.7 percent of the gender-municipality teacher cells did not have a teacher in the young cohort for at least one of the years of the analysis period. The equivalent numbers for the mid-career and old cohorts are 0.7 and 0.2 percent respectively. In the age-specific analyses, I include only gender-municipality-cohorts that have observations for each year of the analysis period.

relative to the mean) and a $6.9 increase in the standard deviation of the monthly wage among mid-career teachers (2.2 percent relative to the mean).[38] The standard deviation of the monthly wage among old teachers is unaffected by the reform. Thus, despite a flattening of the age-wage relationship and an overall reduction in wage dispersion, the reform did fuel statistically and economically significant within-age cohort increases in wage variation.

6.2 The Effect of Wage Decentralization on Teacher Supply

Table 5 displays the coefficient estimates obtained from estimating equation (1) for each of the labor supply outcomes listed in Section 4: age, years of schooling, master's degree, immigrant status, fraction females, fraction on leave, fraction that switched from private school, fraction on temporary contract, average hours worked (measured as a percent of a 40 hour work week), certification status, fraction that move to a different district, fraction that move to a district that had a higher pre-reform non-teacher wage, fraction that remain as teachers, fraction that leave the teaching profession and fraction of new teachers Taken together, the results in Table 5 suggest that the reform changed the composition of teachers toward younger and less formally qualified workers, showing small and statistically significant reductions in age, the fraction of teachers with a master's degree, the fraction with a teaching certificate and average years of education.[39] This is consistent with the wage results presented above, showing a larger wage level effect among young and less experienced teachers. However, none of the effects are large from an economic perspective.[40] For example, the effect of a 1% increase in pre-treatment non-teacher employment income is associated

[38] The identified wage effects may differ depending on the size of the teacher work force in the municipality as well. Larger municipalities face more intense competition over labor, and it is thus possible that local wage-setters in larger municipalities respond more strongly to the reform than do wage-setters in smaller municipalities. In results not shown, I examine this hypothesis by estimating equations (1) and (2) for the municipalities in the bottom 50 percent of the teacher work force size distribution and for the municipalities in the top 50 percent of the teacher work force size distribution, separately. Although the point estimates for the municipalities with larger teacher work forces are larger than those for the municipalities with smaller teacher work forces, the differences in point estimates are not statistically significant.

[39] In results not shown, I also look at the effect of the reform on military entrance exam scores (similar to the AFQT in the US) for the male cohorts that were subject to mandatory military service. I do not find any effects on these outcomes.

[40] In addition to the effects being very small from an economic perspective, event studies for each of these outcomes are relatively noisy, making it difficult to discern any clear effects (Appendix Figures A-2 and A-3).

with a reduction in teacher age of 0.02 years. Given the mean age of teachers shown in Appendix Table A-1 (46.12), this is equivalent to an effect of 0.04% relative to the mean.

The age-specific wage effects presented in Section 4.1 suggest that there may be heterogeneous treatment effects across age cohorts that are not visible in Table 5. This is explored in Appendix Table A-3. These estimates are slightly noisier, but there are no economically and statistically significant differences in the point estimates across the age groups. These results support the idea that the reform did not have a large impact on the composition of teachers from an economic perspective.

It is important to highlight that the substantial size of the teacher workforce means that one would need a very large change in teacher composition to identify an effect in regressions where the dependent variables are based on municipality-gender means. One way to overcome this issue is by repeating the above analysis using only teachers that enter and exit the profession in each year. Any change in the composition of the workforce must be driven by these individuals, and since they constitute only a small fraction of the total workforce, this analysis permits identification of much smaller effects. However, the results obtained from this auxiliary analysis (not shown) are not statistically or economically significantly different from the baseline results.

The findings presented in this section stand in contrast to the results in Biasi (2017), which suggest that pay decentralization leads to an improvement in teacher quality as measured by value-added.[41] However, I investigate the effects of a government-mandated reform, while the analysis in Biasi is based on voluntary adoption of individual bargaining by districts made possible through Wisconsin's Budget Repair Bill. While both settings are interesting, the analyses answer distinctly different questions, and one should not expect the

[41] It should be noted that there is no consensus on the attributes that make a good teacher, and that the teacher characteristics that I examine in this analysis usually are unable to predict value-added (Hanushek and Rivkin 2006; Rockoff et al. 2009; Staiger and Rockoff 2010). The teacher composition results from this analysis are therefore not necessarily at odds with Biasi (2017).

supply effects of these two reforms to be identical. One reason for this is that the Wisconsin Bill changed a number of regulations for all public sector employees in the state (collective bargaining, compensation, health insurance and sick leave). The general equilibrium and wage spillover effects that I discuss in Section 7 may therefore not be present. Another reason is that the Wisconsin Bill affected both teachers and schools on other dimensions than just pay (e.g. retirement benefits for teachers and a cap on spending for schools), and affected the power and structure of teacher collective bargaining. It is possible that a decentralization policy has different effects when accompanied by these changes.[42]

6.3 The Effect of Wage Decentralization on Student Outcomes

Despite the lack of significant teacher composition effects, it is still possible that the reform had an effect on student outcomes. First, the reform and its associated wage effects may impact teacher incentive and motivation in ways that cannot be identified with the outcomes used in Section 6.2. Second, if the wage response to the reform was financed through a reallocation of educational resources, this could also have an effect on student outcomes. To this end, Table 6 displays coefficient estimates obtained from estimating equation (1) for each of the student outcomes listed in Section 4: educational performance in grade 9, the educational performance of these same students at the high school and university level, and the labor market outcomes of these students ten years after graduating from grade 9.[43]

The results in Panel A of Table 6 suggest that the reform did not have a statistically or economically significant effect on educational performance in elementary school as measured by GPA percentile ranking, English percentile ranking and Swedish percentile ranking. While the coefficient on math percentile ranking is marginally statistically significant at the 10 percent level, it is very small from an economic perspective: a 1% increase in pre-treatment

[42] I am unable to examine teacher value-added because students cannot be linked to individual teachers in the registry data.
[43] As I have data from the *LOUISE* registry up to 2012, the cohorts that graduated from 9th grade between 2003 and 2006 have been excluded from the analysis that examine the reform's effect on student labor market outcomes and university enrollment.

college-educated non-teacher employment income is associated with a reduction in math percentile ranking by 0.07% relative to the mean.[44] I can rule out GPA percentile ranking effects larger than 0.07 and smaller than -0.7 from a 10% increase in pre-reform non-teacher employment income.[45]

With respect to high school and university attainment, the results in Table 6 show that the reform had no statistically significant effect on the probability of enrolling in a university-preparatory program or on high school GPA. There is some indication of an adverse effect on the probability of enrolling in a natural science high school track, and of a positive effect on the probability of ever attending university, though these effects are very small. For example, the effect of a 10% increase in pre-treatment non-teacher employment income is associated with a reduction in the probability of enrolling in a natural science high school track of 0.01 and with an increase in the probability of ever being enrolled at university of 0.01.

The results in Panel B of Table 6 further show that the reform did not have an impact on the labor market outcomes of these individuals ten years after graduating from 9th grade; there is no statistically significant effect with respect to employment income, the probability of being in the employment sample, social insurance benefits, government-funded benefits and the probability of being a social security recipient.[46]

The results in this section show that the wage decentralization effect on student short- and long-term education and labor market outcomes was minimal at most. This suggests that unconditional wage decentralization reforms may not generate the efficiency gains implicitly predicted by the existing literature (Britton and Propper 2016; Propper and Van Reenen 2010).

---

[44] Further, event studies for each of these outcomes are relatively noisy, making it difficult to discern any clear effects (Online 4Figures A-5). This also is true with respect to the high school outcomes (Appendix Figure A-5).

[45] The coefficient estimate on GPA is -3.100 with a standard error of 1.942. One can therefore rule out effects larger than $\frac{-3.100+1.96(1.942)}{100} = 0.007$, and smaller than $\frac{-3.100-1.96(1.942)}{100} = -0.07$, from a 1% increase in pre-reform non-teacher wage.

[46] Event studies for university and labor market outcomes are shown in Appendix Figure A-6. These even studies are generally very flat and inconsistent with the idea that the reform had a statistically significant effect on these outcomes.

7. Mechanisms

7.1 Spending and Allocation of Education Resources

The wage effects identified above suggest that pay decentralization leads to an increase in spending on teachers. If this increase is funded through a reduction in spending on other educational inputs that are important for attracting high-quality labor, this may explain the lack of significant supply and productivity effects. This is examined in Figure 5, which depicts results from estimation of equation (2) with respect to (a) per student spending, (b) per student spending on teachers and (c) per student spending on non-teachers. This figure also shows the fraction of total spending dedicated to (d) teachers and (e) non-teachers.

Figure 5 shows that the reform did not have an effect on total education spending (a), but that it did lead to a sizable increase in the amount of resources dedicated toward teachers (b) and to a large reduction in non-teacher spending (c). The long-run response elasticity of teacher spending with respect to non-teacher employment income is 0.2, and the magnitude of this effect is very similar to the wage level effect identified in Section 6. The reform thus led to a reallocation of existing education resources, but not to an overall increase in spending.

To further examine the reform's effect on education spending and resource allocation, Figure 6 shows the effect of the reform on spending on mutually exclusive and collectively exhaustive non-teacher education inputs: (a) food, (b) health, (c) supplies, (d) facilities and (e) uncategorized items. Figure 7 shows the effect of the reform on these same outcomes measured as fractions of total spending.

Despite the relatively large standard errors, subfigures (a) and (e) in Figures 6 and 7 show reductions in total as well as relative spending on food and uncategorized items. While there is a short-term reduction in total and relative spending on facilities (d), this effect disappears in the long-run. With respect to health (b) and supplies (c), the event studies fail to identify significant effects. The event studies in Figures 6 and 7 are consistent with the results in Table 7, obtained from estimation of equation (1) for each of the outcomes.

Overall, Figures 5 through 7 show that the reform did not have an effect on total education spending, but that it did lead to a reallocation of resources away from non-teacher inputs toward teachers. However, these figures provide little evidence to suggest that spending shifted in ways that make teachers and students worse off. Thus, it is unlikely that the reallocation of education resources identified in this section can explain the lack of economically significant teacher composition and student outcome effects.[47]

7.2 General Equilibrium and Wage Spillover Effects

Another explanation for the lack of significant effects on teacher composition and student outcomes is that the reform may have had spillover effects to closely related occupations. Specifically, an increase in teacher wage will make it harder for firms outside the education sector to recruit and retain individuals with teaching degrees. Further, it may induce individuals with non-teaching careers to consider teaching. This will likely put upward pressure on wages in these other industries. My ability to follow individuals across time and space provides a unique opportunity to examine such wage spillover effects.

To examine the presence of wage spillovers, I use the teacher registries discussed in Section 4 to identify all individuals employed as teachers during the two years prior to the reform (1994 and 1995). I merge these data with the *Wage Registry for Public Sector Employees*, which provides me with wage and occupation information for every public sector worker in the country. This allows me to identify which public sector occupations teachers came from, and left for, in the year prior to the reform.[48] As can be seen in Appendix Table A-

[47] There are two other ways through which municipalities can finance an increase in teacher wage: (1) lowering the teacher-student ratio and (2) raising the local income tax rate. However, since education spending was unaffected by the reform it is very unlikely that we would observe an effect on the local income tax rate. In addition, Figure 5 shows that the coefficient on teacher spending closely mirrors the wage effect identified in Section 6, suggesting that the reform did not affected the teacher-student ratio either. Appendix Figure A-7 shows event studies of the reform effect on (a) the local income tax rate and (b) the number of teacher per 100 students, and Appendix Table A-4 shows point estimates obtained from estimation of equation (1) with respect to each of these variables. Appendix Figure A-7 and Table A-4 show that the reform did not have a statistically or economically significant effect on the local tax rate and the teacher-student ratio.

[48] I focus on the pre-reform period to minimize the likelihood that the occupational mobility patterns are confounded by the reform. However, a recent report from Statistics Sweden suggests that these patterns are the same in 2016 as they were prior to the reform (Hellsing 2016). It is important to note that I constrain this part of the analysis to the public sector due to the limitations of

5, more than 80% came from, or left for, jobs within public administration, social services and health care services. Using all workers in the occupations listed in Appendix Table A-5, I estimate equation (2) to examine if the reform affected the wage structure of these occupations, weighting the regressions by the fraction of teachers that came from, and left for, each of the professions.[49] The result from this exercise is shown in Figure 8. The figure reveals the existence of a clear wage spillover effect that starts five years after the reform and stabilizes two years later. This spillover effect thus starts two years after the teacher wage effect but levels out in the same year. The estimated long-run response elasticity is about 0.1, half that of the teacher wage effect.

Table 8 shows the coefficient estimates obtained from estimation of equation (1) for the wage level effects in all substitute occupations listed in Online Table A-5. The results in Table 8 demonstrate that the wage spillover effect is concentrated in the right-tail of the pay distribution in substitute occupations, further supporting the idea that these spillover effects are a consequence of wage-setters in substitute occupations trying to retain and recruit high-quality workers. This underscores the importance of detailed administrative data in evaluations of policies that can induce general equalibrium effects.

The general equilibrium effect identified above coupled with the compressed overall wage distribution in Sweden helps to explain why the reform did not impact the composition of teachers and the long-run outcomes of students (OECD 2001). As can be seen in Appendix Table A-6, the difference in pre-reform non-teacher employment income between municipalities one standard deviation below the mean and one standard deviation above the mean is 7 and 10 percent, for females and males respectively. With a teacher wage response elasticity of 0.2 with respect to outside employment income, the long-run post-reform

---

the *Wage Registry of Private Sector Employees* elaborated on in the data section. However, the majority of teachers in non-teaching occupations work in the public sector, such that this is not a major limitation (Hellsing 2016).

[49] These regressions include gender-by-occupation fixed effects.

difference in absolute teacher wages across 68 percent of the municipalities will thus be less than 2.8 percent for females and 4 percent for males. These numbers are reduced to 1.4 and 2 percent after accounting for the wage spillover effects. It is unlikely that these modest post-reform differences in absolute teacher wages are large enough to overcome standard search and matching frictions and mobility costs.

7.3 Treatment Heterogeneity

A final reason that may help explain why the reform did not have a significant impact on teacher composition and student outcomes relates to treatment heterogeneity. To this end, I follow the existing literature (e.g. Falch 2010) and note that the supply response to wage changes – in particular with respect to mobility - may differ across gender, marital status and parenthood.[50] In results not shown, I have therefore estimated equations (1) and (2) separately for (a) males, (b) females, (c) married teachers, (d) non-married teachers, (e) teachers with at least one child under 18 that lives at home and (f) teachers with no child under 18 that lives at home. The coefficient estimates produced by these stratified regressions are inconsistent with the existence of heterogeneous treatment effects along these dimensions. Consistent with these findings, I find no differential effect of the reform on the wages of these subgroups (Appendix Figure A-8).[51] As the decentralization reform made it easier for local wage-setters to engage in wage discrimination across groups of workers, this is an important finding.

Another potential source of heterogeneity comes from the fact that teachers already earn more than non-teachers in some municipalities, while they earn less than non-teachers in other municipalities. Thus, it is possible that the potential supply effects only are visible in municipalities where the non-teacher wage was higher than the teacher wage prior to the

[50] I have also explored heterogeneity with respect to qualified and unqualified teachers as well as full-time and part-time teachers. However, I do not find any statistically significant differences on these dimensions.
[51] The event studies stratified by gender are shown in Appendix Figure A-1.

reform. However, imposing this sample restriction does not change the coefficient estimates in a statistically significant way (results available upon request).

8. Robustness Checks & Sensitivity Analysis

In this section, I perform a series of sensitivity checks to investigate the robustness of my results to minor alterations of the empirical model. For each of these modifications, I report how the coefficient estimate on teacher wage is affected. Results using the other outcome variables are consistent with these estimates and are available from the author upon request.

The first concern is that there may be persistent transitory fluctuations in earnings in any one year that introduce noise in the treatment measure and attenuate my results (Bhashkar 2005). To this end, I reestimate equation (2) using non-teacher employment income averaged over the five years preceding the reform as the measure of treatment intensity. As illustrated in Appendix Figure A-9 (a), this adjustment has no impact on the economic and statistical significance of the coefficient estimates.

Another concern centers on the possibility that the college-educated non-teacher employment income may not provide an accurate measure of the wage that teachers can command had they not been teachers. An alternative measure can be obtained by first estimating Mincer earnings functions for non-teachers in the year prior to the reform and then using the estimated values from these regressions to predict what the wages of teachers would be had they not been teachers. The result from estimating equation (2) using this alternative measure is depicted in Appendix Figure A-9 (b). The figure shows a teacher wage elasticity that is 0.1 higher than the baseline result.[52] However, the relative time parameter estimates remain within the 95% confidence interval of the baseline result.

[52] In these Mincer earnings functions I control for age, age squared, years of schooling, years of schooling squared, immigrant status, social security recipient status and municipality.

Relatedly, basing the treatment measure on the full distribution of college-educated non-teacher employment incomes may introduce unnecessary noise, as it is unlikely that observations in the tails of the distribution are useful predictors of the wage that teachers could command had they not been teachers. I have therefore also estimated equation (2) excluding gender-specific employment incomes in the top and bottom 5 percent of the distribution. As illustrated in Appendix Figure A-9 (c), this exercise yields a wage effect that is larger than the baseline estimate. However, the relative time parameter estimates remain within the 95% confidence interval of the baseline result.

An issue specific to municipality-level analyses in Sweden during the 90s and 00s stems from the fact that some areas broke away from their municipalities and created their own municipalities during these two decades: Trosa from Nyköping (92), Gnesta from Nyköping (92), Bollebygd from Borås (95), Lekeberg from Örebro (95), Nykvarn from Södertälje (99) and Knivsta from Uppsala (03). Though these newer municipalities are not included in the analysis as only partial time-series information is available, the municipalities that they originally belonged to are. To ensure that the results are robust to excluding these municipalities, as there could be compositional shifts that bias the results, Appendix Figure A-9 (d) shows the result from estimation of equation (2) when these municipalities have been omitted. The results are not different from the baseline results.

An implicit assumption underlying my estimation strategy is that the employment income of college-educated non-teachers is a more accurate reflection of the alternative job market opportunities of teachers than the employment income of non college-educated non-teachers. To examine the validity of this assumption, Appendix Figure A-9 (e) shows how the wage effect of the reform changes when using the pre-reform employment income of non college-educated non-teachers. The figure shows that the baseline result is robust to this adjustment. This is an interesting finding likely driven by the relatively low internal rate of

return to education in Sweden (OECD 2002), a compressed labor market (Kahn 2015), and a strong correlation between college-educated and non college-educated non-teacher employment income within each municipality.[53]

Another assumption behind my estimation strategy is that LLMs, not municipalities, matter when predicting the alternative job market opportunities of teachers. The idea underlying this assumption is that municipalities do not represent unified labor markets and likely fail to fully capture the alternative wage that teachers can command. To examine if my data is consistent with this assumption, Appendix Figure A-9 (f) demonstrates how the wage effect of the reform changes when the treatment measure is based on pre-reform variation in non-teacher employment income at the municipality level. Although the dynamics of the wage response is unaffected, the magnitude of the effect is reduced by approximately 50 percent. This is consistent with the idea that municipalities do not represent unified labor markets and fail to fully capture the alternative wage that teachers can command.

Lastly, a worry specific to the results that examine how the reform affects teacher wages at different deciles of the teacher wage distribution is that it is not clear why the treatment should be based on the average pre-reform non-teacher employment income. To this end, Appendix Table A-7 shows how the wage effect changes depending on which decile of the pre-reform non-teacher employment income distribution that is used to construct the treatment measure. Looking across the rows in Appendix Table A-7, however, it becomes apparent that the wage effect is driven by the general wage level in the municipality, and not by the wage at any one part of the distribution.

[53] During my analysis period, the correlation between college-educated and non college-educated non-teacher employment income within each municipality exceeds 0.9.

9. Discussion & Conclusion

Pre-existing research suggests that centralized wage-setting fuels local labor market distortions that impede market efficiency by forcing local wage-setters to pay wages that deviate from the local competitive equilibrium pay. A natural policy response is to adopt a more flexible wage-setting system that allows wages to better reflect local labor market conditions. Central to this assertion is the idea that wage decentralization can successfully alter the features of the labor market usually associated with centralized wage-setting. However, the lack of exogenous variation in wage-setting regulation has led to a shortage of empirical studies exploring this topic. I address this gap in the literature by exploiting a unique labor market reform in Sweden to present novel evidence on how decentralization affects pay structure, the composition of teachers, and the long-run outcomes of students.

I find that wage decentralization induces considerable changes in teacher pay structure, and that these changes are entirely financed through a reallocation of existing education resources. Specifically, the results reveal a long-run teacher wage response elasticity of 0.2 with respect to the outside wage. The magnitude of the effect is negatively related to teacher age, such that the reform led to a disproportionate increase in entry wage and a flattening of the age-wage relationship. The analysis further shows that the policy fueled an increase in wage dispersion among young and mid-career teachers. I do not find these wage changes to have an economically significant effect on the composition of teachers or on the later-in-life education and labor market outcomes of students.

I show that a main reason for the lack of teacher composition and student outcome effects has to do with wage spillovers to substitute occupations. This effect starts two years after the teacher wage effect and is approximately half as large. The wage spillover effect coupled with the compressed Swedish wage structure helps to explain why the reform did not impact the composition of teachers and the long-run outcomes of students.

In terms of policy implications, the lack of economically significant effects on teacher composition and student outcomes reveals that it may not be possible to reap the predicted benefits of decentralized wage-setting through unconditional decentralization reforms. This could be because municipalities are not able (or willing) to fully eliminate the differences in relative teacher pay across local markets. A solution to this issue would be for the central government to provide financial incentives to local governments to encourage a more substantial wage response. However, my results also suggest that there are large wage spillover effects to non-teaching occupations. Thus, even if the government provides financial incentives to encourage a larger wage response at the local level, it may still not impact teacher composition and student outcomes because nearby industries respond to stay in the competition for labor. The main takeaway from this analysis is thus that the wage decentralization reform induces a reallocation of education resources toward existing teachers. This has no effect on education quality, but leads to a broad cross-industry wage increase due to wage spillovers.

References

Aaronson D., L. Barrow and W. Sander (2007). "Teachers and Student Achievement in Chicago Public High Schools" *Journal of Labor Economics* 25(1): pp. 95-135

Agell, J. (1999). "On the benefits from ridged labour markets: norms, market failures, and social insurance" *Economic Journal* 109(453): pp. 143-64

Agell, J. and K. Lommerud (1992). "Union egalitarianism as income insurance" *Economica* (59): pp. 295-310

Aidt, T. and Z. Tzannatos (2005). "The Cost and Benefits of Collective Bargaining" *Cambridge Working Papers in Economics 0541*

Akerlof, G. (1982). "Labor Contracts as Partial Gift Exchange." *Quarterly Journal of Economics* (97): pp. 543-569

Altonji, J. G. (1988). The effects of family background and school characteristics on educational and labor market outcomes. *Mimeograph* (Northwestern University)

Andréasson, H. (2014). "The effect of decentralized wage bargaining on the structure of firm performance" *The Ratio Institute Working Paper No. 241*

Betts, J. R. (1995). "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth" *The Review of Economics and Statistics* 77(2): pp. 231-250

Bhashkar, M. (2005). "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *Review of Economics and Statistics* 87 (2): pp. 235-255

Biasi, B. (2017). "Unions, Salaries, and the Market for Teachers: Evidence from Wisconsin." Mimeo

Blanchard, O. and L. Summers (1986) "Hysteresis and the European unemployment problem" *NBER Macroeconomic Annual*: pp. 15-77

Booth, A. (1995). The Economics of the Trade Union (Cambridge: Cambridge University Press)

Brewer, D. J. (1996). "Career paths and quit decisions: Evidence from teaching" *Journal of Labor Economics 14*(2): pp. 313–339

Britton J. and C. Propper (2016). "Teacher pay and school productivity: Exploiting wage regulation" *Journal of Public Economics* 133: pp. 75-89

Bruno, M. and J. Sachs (1985). *Economics of Worldwide Stagflation* (Cambridge: Harvard University Press)

Boyd, D., P. Grossman, H. Lankford, S. Loeb and J. Wyckoff (2006). "How changes in entry requirements alter the teacher workforce and affect student achievement" *Education Finance and Policy* 1(2): pp. 176-216

Böhlmark, A. and M. Lindahl (2015). "Independent Schools and Long-run Educational Outcomes: Evidence from Sweden's Large-scale Voucher Reform." *Economica* 82: pp. 508-551

Calmfors, L. and K. Richardson (2004). "Marknadskrafterna och lönebildningen I landsting och regioner" *IFAU Working Paper No. 2004:9*

Calmfors, L. (1993). "Centralization of wage bargaining and macroeconomic performance: a survey" *OECD Economic Studies* (21): pp. 161-191

Calmfors, L. and J. Driffil (1988). "Bargaining structure, corporatism and macro- economic performance" *Economic Policy* 6: pp. 13-62

Cappelli, P. and K. Chauvin (1991). "An interplant test of the efficiency wage hypothesis." *The Quarterly Journal of Economics* 106(3): pp. 769-787

Card, D. (1992). "Using Regional Variation in Wages to Measure the Effects of the Federal Minimum Wage." *ILR Review* 46(1): pp. 22-37

Card, D. and S. de la Rica (2006). "Firm-Level Contracting and the Structure of Wages" *Industrial and Labor Relations Review* (59): pp. 573-593

Cardoso, A. and P. Portugal (2005). "Contractual Wages and the Wage Cushion under Different Bargaining Settings" *Journal of Labor Economics 23(4):* pp. 875-902

Cardullo, G. (2015). "The Welfare and Employment Effects of Centralized Public Sector Wage Bargaining" *MPRA Working Paper No. 66879*

Census (1996). "Money Income in the United States: 1995" in *Current Population Reports: Consumer Income* (Washington, D.C.: U.S. Government Printing Office)

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

Clotfelter, C., E. Glennie, H. Ladd and J. Vigdor (2008). "Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina" *Journal of Public Economics* 92: pp. 1352-1370

Clotfelter, C., H. Ladd and J. Vigdor (2007) "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects" *CALDER Working Paper 11* (Washington, DC: The Urban Institute)

Corcoran, S., W. Evans and R. Schwab (2004). "Women, the labor market, and the declining relative quality of teachers" *Journal of Policy Analysis and Management* 23: pp. 449-470

Crouch, C. (1985). "Conditions for trade union wage restraint" in L. Lindberg and C. Maier (eds.), *The politics of inflation and economic stagflation: theoretical approaches and international case studies* (Washington, D.C.: The Brookings Institution)

Dahl, C. M., D. le Maire and J. R. Munch (2013). "Wage Dispersion and Decentralization of Wage Bargaining" *Journal of Labor Economics* 31(3): pp. 501-533

Daouli, J., M. Demoussis, N. Giannakopoulos and I. Laliotis (2013). "Firm-Level Collective Bargaining and Wages in Greece: A Quantile Decomposition Analysis" *British Journal of Industrial Relations 51(1)*: pp. 80-103

Darling-Hammond, L. (2000). "Teacher Quality and Student Achievement: A Review of State Policy Evidence" *Education Policy Analysis Archives* 8(1): pp. 1-44

Dell'Aringa, C. and C. Lucifora (1994a). "Collective Bargaining and Relative Earnings in Italy" *European Journal of Political Economy 10*: pp. 727-747

Dell'Aringa, C. and C. Lucifora (1994b). "Wage Dispersion and Unionism: Do Unions Protect Low Pay?" *International Journal of Manpower 15(2/3):* pp. 150-170

Dell'Aringa, C. and L. Pagani (2007). "Collective Bargaining and Wage Dispersion in Europe" *British Journal of Industrial Relations* (45): pp. 29-45

Dolton, P. and O. Marcenaro-Guiterrez (2011). "If you pay peanuts do you get monkeys? A cross-country analysis of teacher pay and pupil performance" *Economic Policy* 26(65): pp. 5-55

Edin, P.A. and B. Holmlund (1995). The Swedish Wage Structure: The Rise and Fall of Solidarity Wage Policy? In In R. Freeman and L. Katz (eds.), *Differences and Changes in Wage Structures* Chicago: University of Chicago Press

Falch, T. (2010). "The Elasticity of Labor Supply at the Establishment Level" *Journal of Labor Economics* 28(2): pp. 237-266

Falch, T. and B. Strom (2006). "Local Flexibility in Wage Setting: Evidence From the Norwegian Local Public Sector" *Empirical Economics* 31(1): pp. 113-142

Figlio, D. (2002). "Can public schools buy better-qualified teachers?" *Industrial and Labor Relations Review 55*(4): pp. 686–697

Fitzenberger, B., K. Kohn and A.C. Lembcke (2008). "Union Density and Varieties of Coverage: The Anatomy of Union Wage Effects in Germany" *IZA Discussion Paper* (3356)

Fitzberger, B., K. Kohn and A. Lembcke (2013). "Union Density and Varieties of Coverage: The Anatomy of Union Wage Effects in Germany" *ILR Review 66(1):* pp. 169-197

Freeman, R. and R. Gibson (1995). "Getting together and breaking apart: the decline of centralized collective bargaining" in R. Freeman and L. Katz (eds.), *Differences and Changes in Wage Structures* (Chicago: University of Chicago Press)

Gerlach, K. and S. G. Stephan (2005). "Wage Distributions by Wage-setting Regime" *IAB Discussion Papper No. 200509*

Goldhaber, D. (2002).' The Mystery of Good Teaching.' *Education Next*, vol. 2(1)

Granqvist, L. and H. Regnér (2008). "Decentralized Wage Formation in Sweden" *British Journal of Industrial Relations*: pp. 500-520.

Gritz, R., & Theobald, N. (1996). "The effects of school district spending priorities on length of stay in teaching" *Journal of Human Resources 31*(3): pp. 477–512

Grogger, J. (1996). "Does School Quality Explain the Recent Black/White Wage Trend?" *Journal of Labor Economics*

Guarino, C., L. Santibanez and G. A. Daley (2006). "Teacher Recruitment and Retention: A Review of the Recent Empirical Literature" *Review of Educational Research* 76(2): pp. 173-208

Gurtzgen, N. (2007). "The Effect of Firm- and Industry-Level Contracts on Wages – Evidence from Longitudinal Linked Employer-Employee Data" *Discussion Paper 06-82* (ZEW, Mannheim)

Harcourt, G. (1997). "Pay policy, accumulation and productivity" *Economic and Labour Relations Review* 8: pp. 78-89

Hansson, R. (2015). "Större andel kvinnliga lärare I grundskolan" *SCB No. 2015:149*. Accessed January 12, 2017, from: http://www.scb.se/sv_/Hitta-statistik/Artiklar/Storre-andel-kvinnliga-larare-i-grundskolan/

Hanushek, E., J. Kain and S. Rivkin (2004). "Why public schools lose teachers" *Journal of Human Resources 39*(2): pp. 326–354

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

Hanushek, E., Kain, J. and Rivkin, S. (1998). Teachers, Schools and Academic Achievement. *Econometrica*

Hanushek E. (2003). "The failure of input-based schooling policies" *The Economic Journal* 113: pp. 64-98

Hanushek, E. and S. Rivkin (2007). "Pay, Working Conditions, and Teacher Quality" *Future of Children* 17(1): pp. 69-96

Heckman, J., L. Lochner and P. Todd (2006). "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond" in *Handbook of the Economics of Education* (eds. E. Hanushek and F. Welch)

Hellsing, Eric (2016). "40000 lärare arbetar inte med undervisning" *Statistics Sweden No. 2016:4*

Hendricks, M. (2014). "Does it pay to pay teachers more? Evidence from Oklahoma's minimum salary schedule" *Journal of Public Economics* (109): pp. 50-63

Hibbs, D. and H. Locking (1995). "Den solidariska lönepolitiken och produktiviteten inom industrin" *Ekonomisk Debatt* 23(7): pp. 537-548

Hoxby, C. and A. Leigh (2004). "Pulled Away or Pushed Out? Explaining the Decline of Teacher Aptitude in the US" *American Economic Review* 94(2): pp. 236-240

Ibsen, C., T. Larsen, J. Madsen and J. Due (2011). "Challenging Scandinavian employment relations: the effects of new public management reforms" *International Journal of Human Resource Management 22(11):* pp. 2295-2310

Imazeki, J. (2005). "Teacher Salaries and Teacher Attrition" *Economics of Education Review* 24: pp. 431-449

Iversen, T. (1996). "Power, Flexibility, and the Breakdown of Centralized Wage Bargaining: Denmark and Sweden in Comparative Perspective." *Comparative Politics* 28(4): pp. 399-436

Ingersoll, R. and H. May (2012). "The magnitude, destinations, and determinants of mathematics and science teacher turnover" *Educational Evaluation and Policy Analysis* 34(4): pp. 435-464

Kahn, L. (1998). "Collective Bargaining and the Interindustry Wage Structure: International Evidence." *Economica* 65(260): pp. 507-534

Kahn, L. (2015). "Wage Compression and the Gender Pay Gap" *IZA World of Labor 150*

Karlsson, N. and H. Lindberg (2011). "The Decentralization of Wage Barganing: Four Cases" *The Ratio Institute Working Paper No. 178*

Karlsson, N. and H. Lindberg (2008). "En ny svensk modell – Vägval på arbetsmarknaden: Sönderfall, omreglering, avreglering eller modernisering*" The Ratio Institute*

Kirby, S., Berends, M., & Naftel, S. (1999). "Supply and demand of minority teachers in Texas: Problems and prospects" *Educational Evaluation and Policy Analysis 21*(1): pp. 47–66

Lankford, M., S. Loeb & J. Wyckoff (2002). "Teacher sorting and the plight of urban schools: A descriptive analysis" *Educational Evaluation and Policy Analysis 24*(1): pp. 37–62

Layard, R., S. Nickell and R. Jackman (1991). Unemployment (Oxford: Oxford University Press)

Leigh, A. (2012). "Teacher pay and teacher aptitude" *Economics of Education Review 31*: pp. 41-53

Loeb, S. and M. E. Page (2000). "Examining the link between teacher wages and student outcomes: The importance of alternative labor market opportunities and non-pecuniary variation" *The Review of Economics and Statistics* 82(3): pp. 393-408

Lärarförbundet (2016). "Vi Behöver Fler Lärare". Accessed on July 6, 2017, from: https://www.lararforbundet.se/artiklar/lararbrist-sverige-behover-fler-larare

Mincer, J. (1974). *Schooling, Experience and Earnings* (New York: Columbia University Press)

Moene, K. and M. Wallerstein (1993). "The Economic performance of different bargaining institutions: a survey of the theoretical literature" *Wirtchaft und Gesellschaft* (19): pp. 423-450

NCES (2013). *Characteristics of Public and Private Elementary and Secondary School Teachers in the United States: Results From the 2011-12 Schools and Staffing Survey.* January 12, 2017, from: http://nces.ed.gov/pubs2013/2013314.pdf

OECD (1996). *Teachers' pay and conditions*. Retrieved February 16, 2016, from: http://www.oecd.org/dataoecd/39/62/1840245.pdf

OECD (2001). *Divided We Stand: Why Inequality Keeps Rising* (Paris: OECD)

OECD (2002). *Education at a Glance* (Paris: OECD)

OECD (2004). *Employment Outlook* (Paris: OECD)

OECD (2015). *Teachers Matter: Attracting, Developing and Retaining Effective Teachers* (Paris: OECD)

OECD (2016), Teachers' salaries. doi: 10.1787/f689fb91-en (Accessed on 07 July 2016)

Plasman, R., M. Rusinek and F. Rycx (2007). "How Do Company Collective Agreements Affect Wages? Evidence from Four Corporatist Countries" *European Journal of Industrial Relations* 13(2): pp. 161-180

Podgursky, M., R. Monroe and D. Watson (2004). "The academic quality of public school teachers: An analysis of entry and exit behavior" *Economics of Education Review* 23: pp. 507–518

Propper, C., and J. van Reenen (2010). "Can pay regulation kill? Panel data evidence on the effect of labor markets on hospital performance" *Journal of Political Economy* 118(2)

Rivkin, S., E. Hanushek and J. Kain (2005). "Teachers, Schools, and Academic Achievement" *Econometrica* (73): pp. 417-58

Rockoff, J. (2004). "The impact of individual teachers on student achievement: Evidence from Panel Data" *The American Economic Review* 94(2): pp. 247-252

Ronfeldt, M., H. Lankford, S. Loeb and J. Wyckoff (2011). "How Teacher Turnover Harms Student Achievement" *American Educational Research Journal* 50(1): pp. 4-36

Roy, A.D. (1951). "Some Thoughts on the Distribution of Earnings" *Oxford Economic Papers* (3): pp. 135-146

Rycx, F. (2003). "Industry Wage Differentials and the Bargaining Regime in a Corporatist Country" *International Journal of Manpower* (24): pp. 347-366

Sapsford, D. and Z. Tzannatos (1993) The Economics of the Labour Market (London: MacMillan Publishing)

Shapiro, C. and J. Stiglitz (1984). "Equilibrium unemployment as a worker discipline device." *American Economic Review* (74): pp. 433-44

Skolverket (2014). *Nästan alla grundskoleelever fortsätter till gumnasieskolan*. Accessed December 27, 2016, from: https://www.skolverket.se/statistik-och-utvardering/nyhetsarkiv/nyheter-2014/nastan-alla-grundskoleelever-fortsatter-till-gymnasieskolan-1.223182

Staiger, D. O. and J. E. Rockoff (2010). Searching for effective teacher with imperfect information. *Journal of Economic Perspectives* 24(3): pp. 97-118

Statistics Sweden (2006). *Uppgifter på kommunnivå, Tabell 1: Skolor, elever och språkval läsåret 2006/07*. Accessed February 14, 2017, from: http://www.skolverket.se/statistik-och-utvardering/statistik-i-tabeller/grundskola/skolor-och-elever/skolor-och-elever-i-grundskolan-lasar-2006-07-1.39770

Slater, H., N. Davis and S. Burgees (2012). "Do teachers matter? Measuring the variation in teacher effectiveness in England" *Oxford Bulletin of Economics and Statistics* 74(5): pp. 629-645

Stockard, J., & Lehman, M. (2004). "Influences on the satisfaction and retention of 1st-year teachers: The importance of effective school management" *Educational Administration Quarterly 40*(5): pp. 742–771

Sockice, D. (1990). "Wage determination: The changing role of institutions in advanced industrialized countries" *Oxford Review of Economic Policy* (6): pp. 36-61

Summers, L., J. Gruber and R. Vergara (1993). "Taxation and the Structure of the Labor Market: the Case of Corporatism." *Quarterly Journal of Economics* 108: pp. 384-411

Söderström, M. (2006). "Evaluating Institutional Changes in Education and Wage Policy" *IFAU Dissertation Series 2006:3*.

Traxler, F. (2003). "Bargaining (De)centralization, Macroeconomic Performance and Control over the Employment Relationship" *British Journal of Industrial Relations 41(1):* pp. 1-27

Weiss, E. (1999). "Perceived workplace conditions and first-year teachers' morale, career choice commitment, and planned retention: A secondary analysis" *Teaching and Teacher Education 15*(8): pp. 861–879

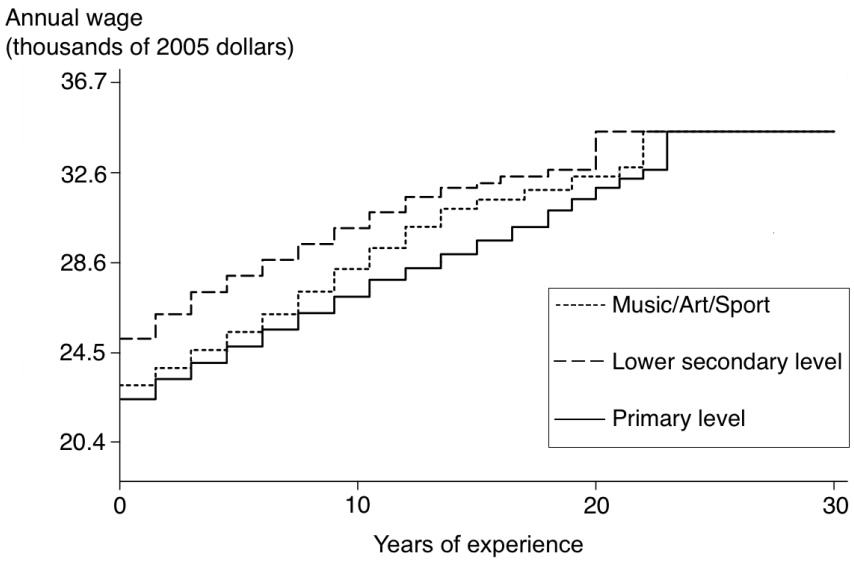


Figure 1.1: 1990 steps-and-lanes salary schedule at the elementary school level (in 000's)

Notes: The wage information depicted in this figure is taken from Söderström (2006). See Section III for a detailed description of the teacher pay scales that were used prior to the 1996 decentralization reform.

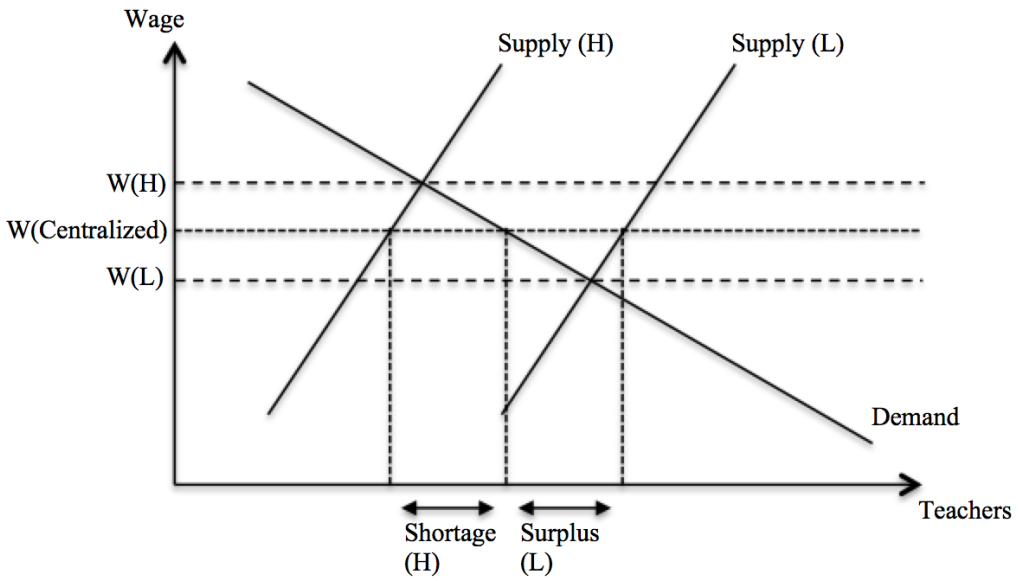
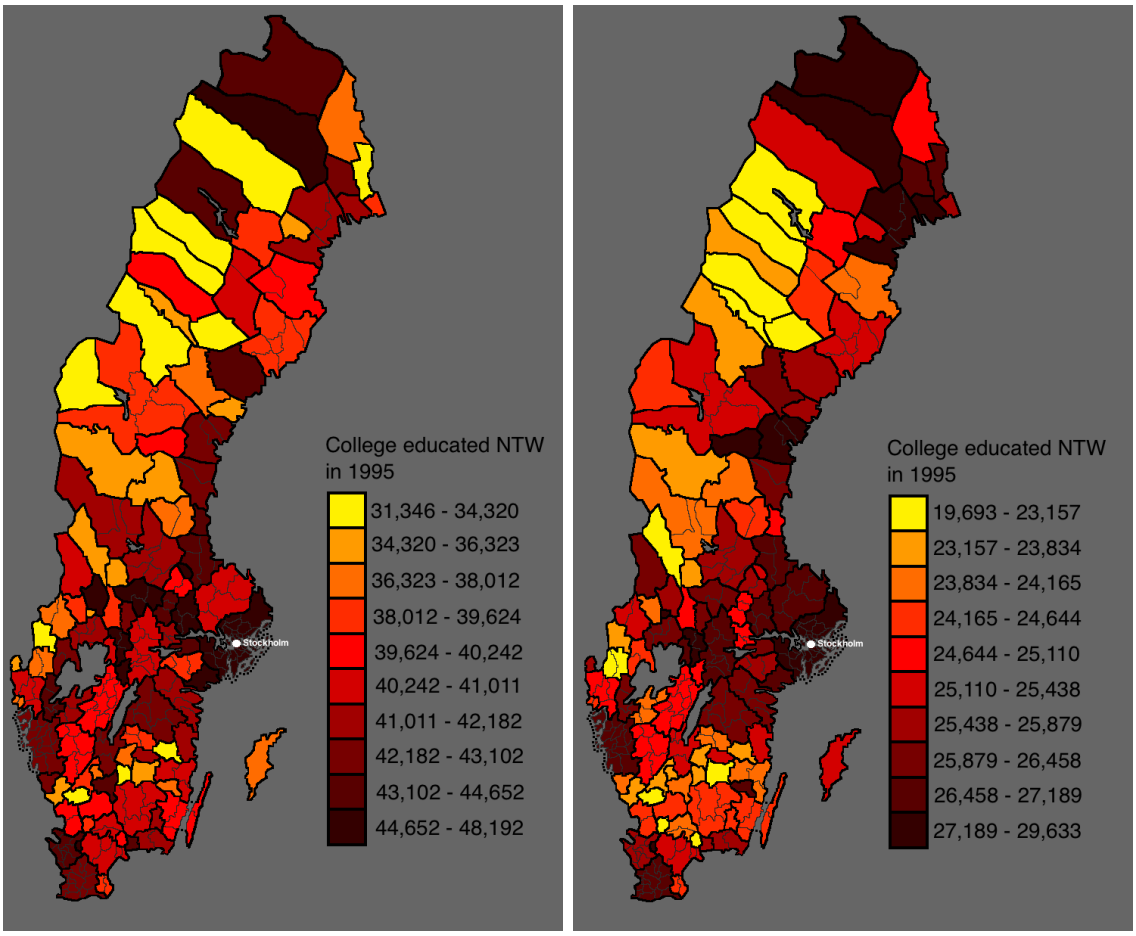


Figure 1.2: Centralized wage-setting

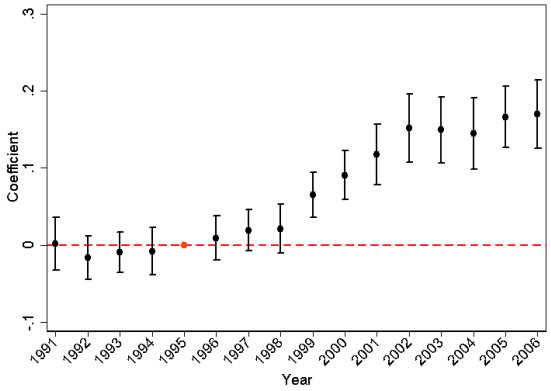
Notes: Figure based on Britton and Propper (2016). L represents the low-productivity region and H depicts the high-productivity region. See Section III for a detailed description of the anticipated labor market implications associated with centralized wage-setting.



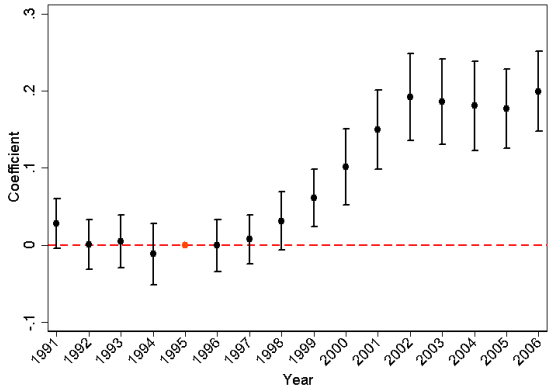
(a) Males **(b)** Females

Figure 1.3: Geographic variation in pre-reform gender-specific college-educated non-teacher employment income across local labor markets

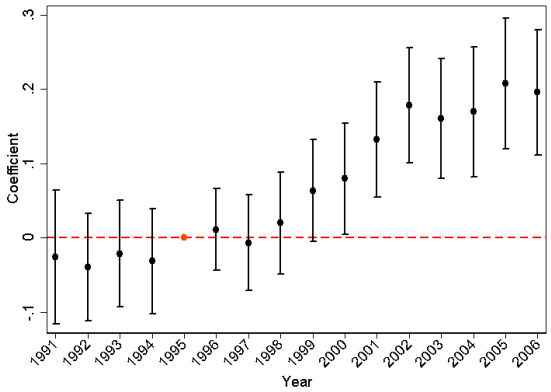
Notes: The maps show geographic variation in college-educated non-teacher employment income (thousands of 2005 dollars) across local labor markets in 1995 for males (a) and females (b) respectively. The gender-specific employment income distributions have been divided into deciles, with local labor markets in yellow belonging to the bottom decile and local labor markets in brown belonging to the top decile. Black solid lines indicate 1995 local labor market boarders while gray solid lines indicate 1995 municipality boarders. Each local labor market boarder is also a municipality boarder. The dotted black lines are used to show that all islands inside those boarder also belong to the local labor market.



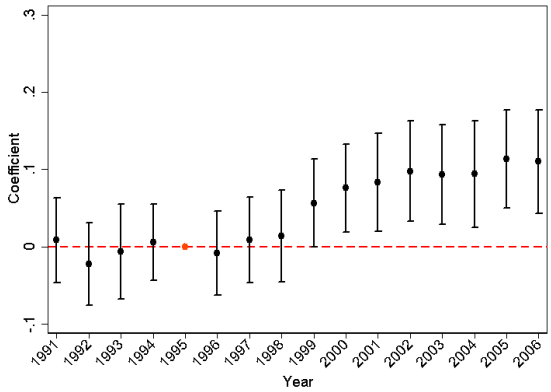
(a) Mean



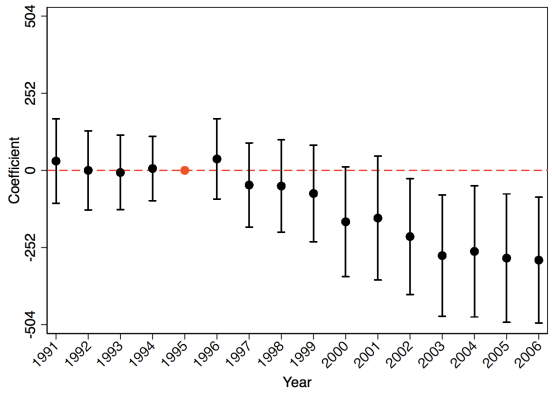
(b) Median



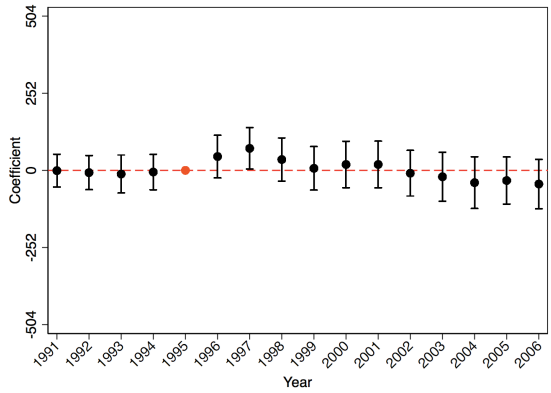
(c) 10th Percentile



(d) 90th Percentile



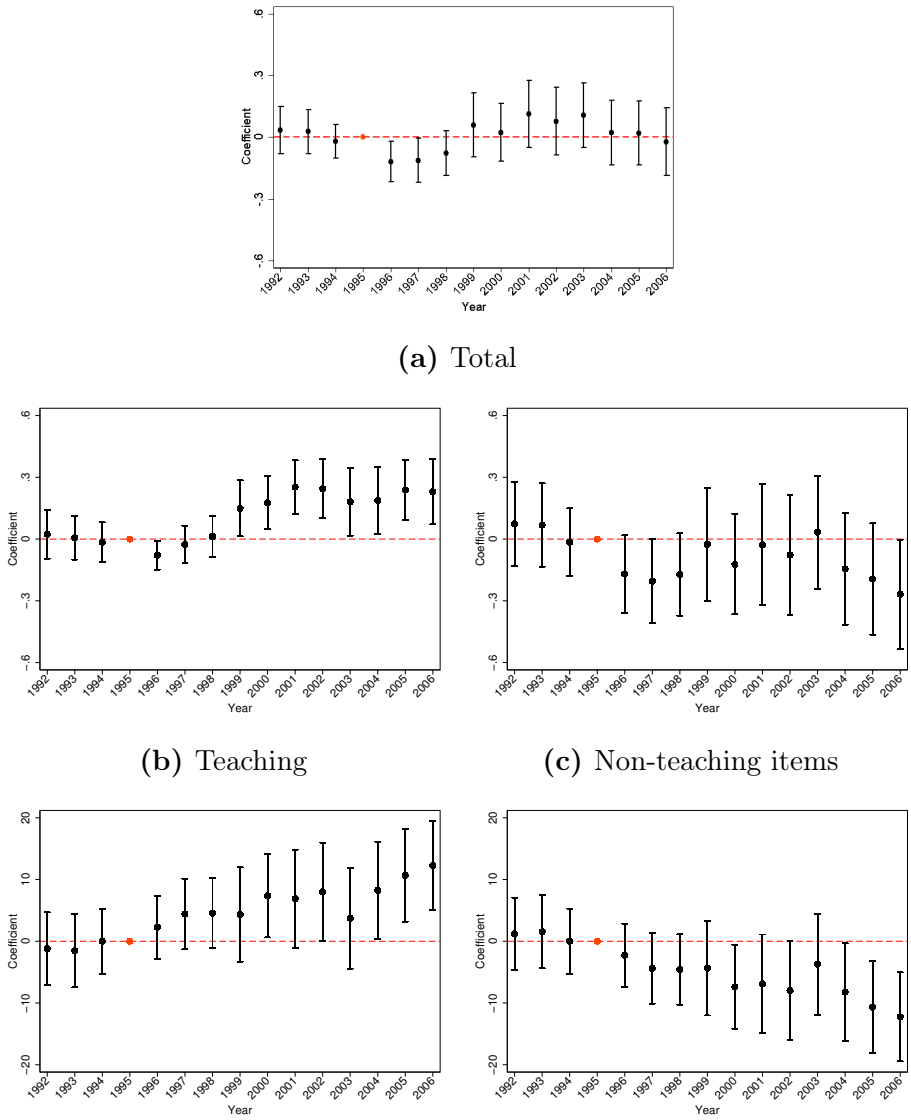
(e) Interquartile range



(f) Standard deviation

Figure 1.4: Event study estimates - wage structure

Notes: Author's estimation of equation (2) as described in the text using 1991-2006 teacher registry data on all public elementary school teachers in Sweden. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.



(a) Total

(b) Teaching

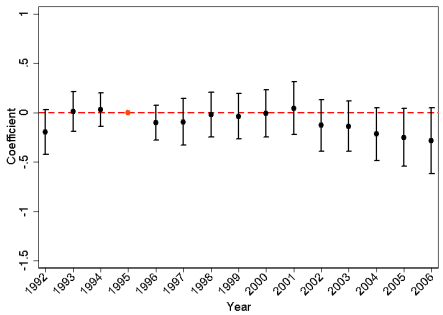
(c) Non-teaching items

(d) Fraction Teaching

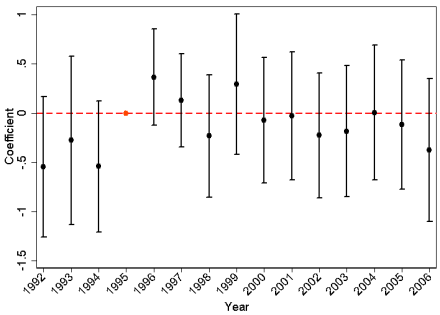
(e) Fraction Non-Teaching

Figure 1.5: Event study estimates - Per student education spending

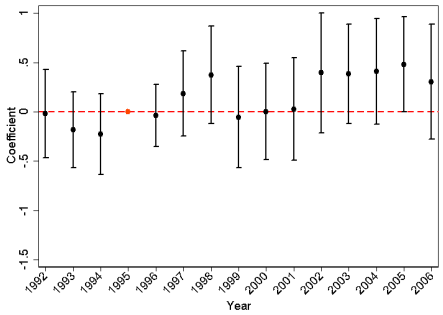
Notes: Author's estimation of equation (2) as described in the text using 1992-2006 registry data supplemented with public-use data from Statistics Sweden and SNAE. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.



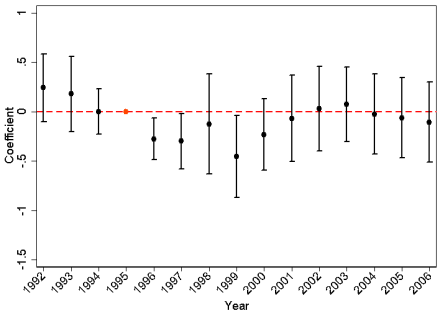
(a) Food



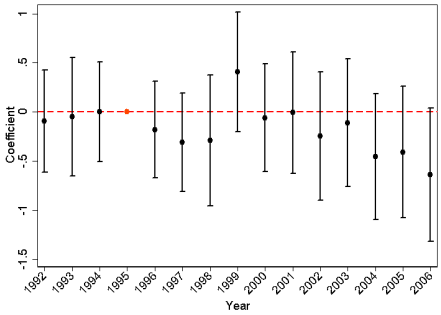
(b) Health



(c) Supplies



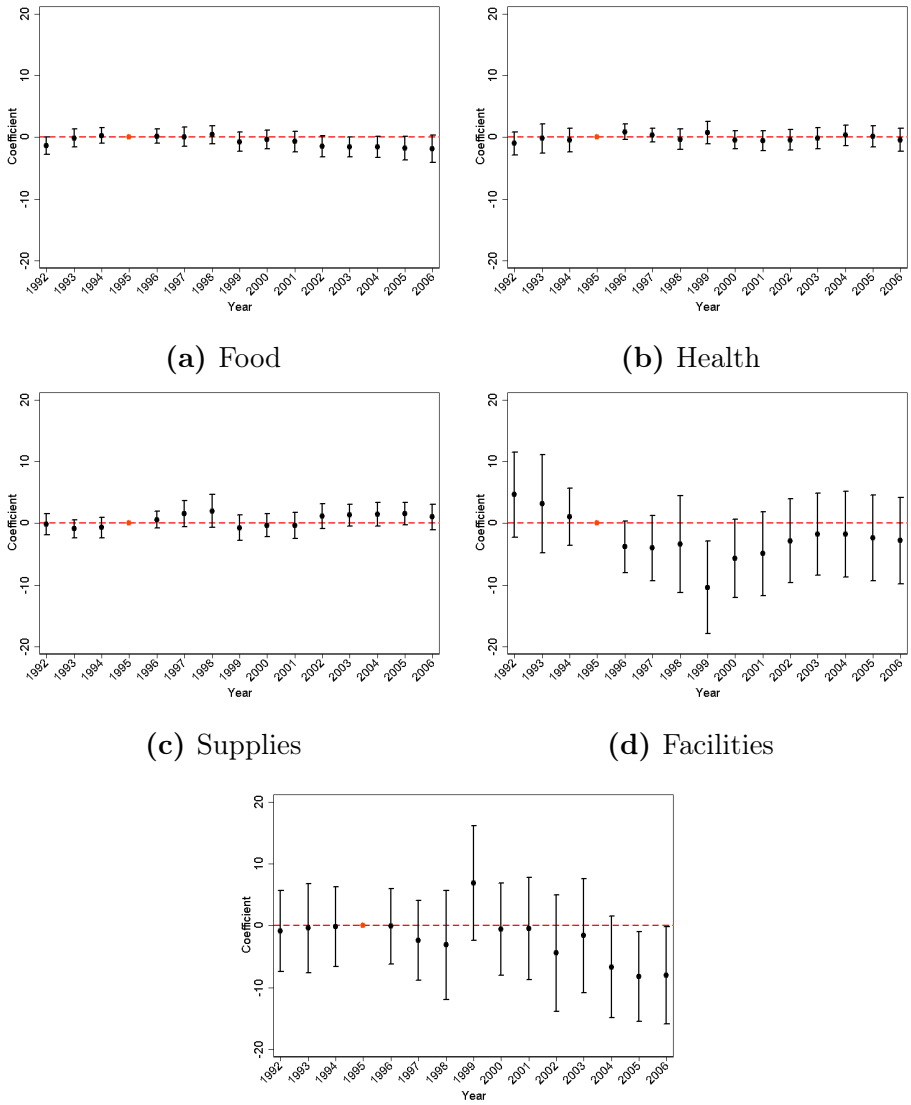
(d) Facilities



(e) Uncategorized items

Figure 1.6: Event study estimates - Per student education spending

Notes: Author's estimation of equation (2) as described in the text using 1992-2006 registry data supplemented with public-use data from Statistics Sweden and SNAE. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.



(a) Food

(b) Health

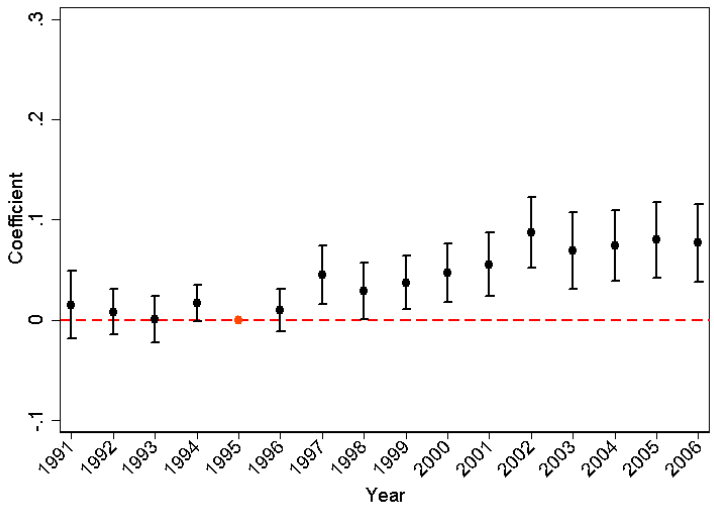
(c) Supplies

(d) Facilities

(e) Uncategorized items

Figure 1.7: Event study estimates - Spending on education inputs as a fraction of total education spending

Notes: Author's estimation of equation (2) as described in the text using 1992-2006 registry data supplemented with public-use data from Statistics Sweden and SNAE. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.



(a) All substitute occupations

Figure 1.8: Event study estimates - Wage spillover effect

Notes: The unit of observation is a municipality-gender-occupation. The sample includes all workers in the three-digit public occupation groups that any teacher switched to/from in the year prior to the reform. The estimates include municipality-gender, year-gender and occupation-gender fixed effects, as well as controls for gender composition, average years of schooling, fraction of social security benefit recepients, mean income, average age and fraction immigrants. Regressions are weighted by the fraction of teachers that switched to each of the occupation groups. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.

Table 1.1: Summary statistics of public elementary school teachers and non-teachers

	Full Period		1991-1995		1996-2000		2001-2006	
	Teachers	Non-teachers	Teachers	Non-teachers	Teachers	Non-teachers	Teachers	Non-teachers
Male	0.269	0.512	0.285	0.511	0.266	0.512	0.259	0.512
Age	45.433	41.299	45.578	40.436	45.277	41.221	45.440	42.044
Immigrant	0.090	0.148	0.077	0.126	0.086	0.146	0.104	0.168
Married	0.636	0.495	0.701	0.537	0.632	0.495	0.586	0.462
Partners	0.070	0.096	0.046	0.081	0.063	0.095	0.094	0.108
Child	0.593	0.510	0.646	0.533	0.579	0.508	0.561	0.495
Single parent	0.077	0.089	0.076	0.084	0.076	0.088	0.079	0.093
Years of schooling	14.941	11.612	14.986	11.297	14.916	11.547	14.923	11.913
Masters degree	0.262	0.053	0.276	0.041	0.248	0.048	0.263	0.067
Social security recipient	0.013	0.059	0.017	0.076	0.017	0.077	0.005	0.031
Observations	1,318,287	82,194,429	407,236	24,898,410	398,706	25,705,147	512,345	31,590,872

Notes: Author's calculation using 1991-2006 teacher registry data on all public elementary school teachers in Sweden and 1991-2006 LOUISE registry data on all non-teachers in Sweden that are over 18 years old but less than 66 years old.

Table 1.2: Pre-reform dependent variable differences in means for public elementary school teachers

	Bottom 10 Percent	Top 10 Percent	Difference in Means	
	Mean	Mean	Difference	T statistic
Years of Schooling	14.923	14.672	0.251	6.104***
Age	45.325	46.156	-0.831	-2.483**
Temporary Contract	0.123	0.158	-0.035	-3.059***
Percent Work	91.356	92.616	-1.259	-2.386**
Immigrant	0.046	0.105	-0.060	-5.486***
Female	0.695	0.737	-0.042	-4.423***
Mover	0.011	0.008	0.003	0.931
On Leave	0.015	0.024	-0.009	-1.707*
Private Switch	0.001	0.003	-0.002	-1.916*
Certificate	0.922	0.852	0.069	6.744***
Stayers	0.850	0.822	0.027	1.932*
Leavers	0.150	0.178	-0.027	-1,932*
Hires	0.138	0.161	-0.023	-1.809*

Notes: Author's calculation using 1995 teacher registry data on all public elementary school teachers in Sweden. "Bottom 10 Percent" refers to the municipalities at the bottom decile of the college-educated non-teacher employment income distribution, while "Top 10 Percent" refers to the municipalities at the top decile of the college-educated non-teacher employment income distribution. Column 6 depicts the differences in means between the two groups, and is equal to the value in Column 2 minus the value in Column 4. The T-statistic shows the Student t-statistic associated with the null hypothesis that the differences in means between the two groups are zero, allowing for differences in variances across the two groups. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level. These stars are based on two-tailed p-values.

Table 1.3: The effect of wage decentralization on pay structure

	Mean	Median	10th percentile	90th percentile	IQR	SD
lnNTW*Post	0.104***	0.109***	0.130***	0.068***	-161.837**	11.087
	(0.015)	(0.017)	(0.029)	(0.017)	(63.676)	(26.332)

Notes: Author's estimation of equations (1) as described in the text using 1991-2006 teacher registry data on all public elementary school teachers in Sweden. Regressions are based on 9120 municipality-gender observations. All estimates include municipality-gender and year-gender fixed effects, as well as controls for gender composition, average years of schooling, fraction of social security benefit recepients, mean income, average age and fraction immigrants. Standard errors clustered at the municipality level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 1.4: The effect of wage decentralization on pay structure stratified by teacher age

	Wage level				Wage dispersion	
	Mean Wage	Median Wage	10th Percentile Wage	90th Percentile Wage	Interquartile Range	Standard Deviation
Panel A: Teachers 20 to 34 Years Old						
lnNTW*Post	0.161***	0.088**	0.200***	0.174***	137.659**	92.894**
	(0.028)	(0.040)	(0.034)	(0.030)	(66.272)	(36.059)
Panel B: Teachers 35 to 49 Years Old						
lnNTW*Post	0.134***	0.117***	0.140***	0.132***	8.974	69.385***
	(0.018)	(0.036)	(0.018)	(0.020)	(45.471)	(25.072)
Panel B: Teachers 50 to 64 Years Old						
lnNTW*Post	0.080***	0.159***	0.107***	0.031*	26.647	7.631
	(0.022)	(0.042)	(0.026)	(0.018)	(51.266)	(36.311)

Notes: Author's estimation of equations (1) as described in the text using 1991-2006 teacher registry data on all public elementary school teachers in Sweden. Regressions are based on 9120 municipality-gender observations. All estimates include municipality-gender and year-gender fixed effects, as well as controls for gender composition, average years of schooling, fraction of social security benefit recepients, mean income, average age and fraction immigrants. Standard errors clustered at the municipality level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 1.5: The effect of wage decentralization on the composition of teachers

	Age	Master's Degree	Immigrants	Female	On leave	Switched from private school	Temporary contract	Percent worked
lnNTW*Post	-1.980*	-0.068**	-0.007	-0.054***	-0.004	0.010***	0.028	0.005
	(1.116)	(0.030)	(0.016)	(0.015)	(0.009)	(0.003)	(0.036)	(0.008)
% Effect	-0.043	-0.267	-0.093	-0.074	-0.182	5.000	0.174	0.005
	Certificate	Movers	High Movers	Stayers	Leavers	Hires	Years of Schooling	
lnNTW*Post	-0.087***	0.002	0.130	0.012	-0.012	0.038	-0.313***	
	(0.030)	(0.006)	(0.125)	(0.028)	(0.028)	(0.027)	(0.112)	
% Effect	-0.101	-0.066	0.182	0.278	0.014	-0.076	0.262	

Notes: Author's estimation of equations (1) as described in the text using 1991-2006 teacher registry data on all public elementary school teachers in Sweden supplemented with 1992-2006 public-use data from Statistics Sweden and SNAE. Regressions are based on 8550 municipality-gender observations. All estimates include municipality-gender and year-gender fixed effects, as well as controls for gender composition, average years of schooling, fraction of social security benefit recepients, mean income, average age, fraction immigrants, fraction of student that attend private schools and total number of elementary school students. Standard errors clustered at the municipality level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 1.6: The effect of wage decentralization on student education and labor market outcomes

Panel A: 9th Grade & High School							
	9th Grade Outcomes				High School Outcomes		
	GPA	Swedish	English	Math	University-Prep. HS Program	Natural Science HS Track	HS GPA
lnNTW*Post	-3.100	-2.539	-2.606	-3.719*	-0.034	-0.064**	0.547
	(1.942)	(1.735)	(1.942)	(1.975)	(0.043)	(0.028)	(2.658)
% Effect	-0.063	-0.052	-0.054	-0.076	-0.069	-0.356	0.011
Panel B: Labor Market & Higher Education							
	Employment Income	Employment Sample	Social Security Recipient	Social Insurance Benefits	Government-Funded Benefits	University Enrollment	
lnNTW*Post	0.026	0.019	-0.005	0.045	-0.073	0.065***	
	(0.059)	(0.019)	(0.009)	(0.069)	(0.127)	(0.022)	
% Effect	-	0.021	0.109	-	-	0.243	

Notes: Author's estimation of equations (1) as described in the text using 1992-2006 registry data supplemented with public-use data from Statistics Sweden and SNAE. Regressions in Panel A are based on 8550 municipality-gender-year observations. The results in Panel B are based on 6270 observations as these outcomes are measured 10 years after students graduate from 9th grade, and the most recent labor market data I have access to is from 2012. Students that graduated 9th grade between 2003 and 2006 are therefore excluded from Panel B. All estimates include municipality-gender and year-gender fixed effects, as well as controls for gender composition, average years of schooling, fraction of social security benefit recepients, mean income, average age and fraction immigrants. Standard errors clustered at the municipality level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 1.7: The effect of wage decentralization on per student education spending and resource allocation

	Total	Non-teaching	Facilities	Teaching	Supplies	Health	Food	Other
Panel A: Per student spending								
lnNTW*Post	-0.001	-0.119**	-0.265*	0.118**	0.311	0.326	-0.065	-0.154
	(0.005)	(0.047)	(0.148)	(0.050)	(0.206)	(0.219)	(0.086)	(0.200)
Panel B: Resource allocation								
lnNTW*Post		-6.799***	-6.347**	6.799***	1.264*	0.457	-0.404	-1.769
		(2.286)	(2.584)	(2.286)	(0.734)	(0.642)	(0.560)	(2.402)

Notes: Author's estimation of equations (1) as described in the text using 1992-2006 registry data supplemented with public-use data from Statistics Sweden and SNAE. Dependent variables in Panel B are the percent of total per student public elementary education spending dedicated to that particular input. Regressions are based on 8550 municipality-gender observations. All estimates include municipality-gender and year-gender fixed effects, as well as controls for gender composition, average years of schooling, fraction of social security benefit recepients, mean income, average age, fraction immigrants, fraction of student that attend private schools and total number of elementary school students. Standard errors clustered at the municipality level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 1.8: The effect of teacher wage decentralization on pay in substitute occupations

	Mean	Median	10th percentile	90th percentile
lnNTW*Post	0.047***	0.044***	0.023	0.075***
	(0.012)	(0.012)	(0.016)	(0.025)

Notes: Author's estimation of equations (1) as described in the text using 1991-2006 registry data. The unit of observation is a municipality-gender-occupation. Panel A includes all three-digit public occupation groups that any teacher switched to/from in the year prior to the reform. Panel B includes the three most popular three-digit public occupation groups that teachers switched to/from in the year prior to the reform (public administration, social services and health care services). All estimates include municipality-gender, year-gender and occupation-gender fixed effects, as well as controls for gender composition, average years of schooling, fraction of social security benefit recepients, mean income, average age and fraction immigrants. Regressions are weighted by the fraction of teachers that switched to each of the occupation groups. Standard errors clustered at the municipality level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

CHAPTER 2

Tipping and the Effects of Segregation[a]

Abstract

Card, Mas and Rothstein (2008) show that residential segregation can arise due to white aversion toward minority neighbors: when a neighborhood's minority share passes a certain "tipping point," the neighborhood experiences white flight and avoidance. We demonstrate that the dynamics of segregation between immigrants and natives in Sweden's metropolitan areas also is characterized by this phenomenon. We then add to the neighborhood effects literature by using the tipping threshold as a source of exogenous variation in ethnic population composition to provide evidence on how neighborhood segregation affects individual outcomes. We find negative effects on the educational attainment of native children. These effects are temporary and do not carry over to the labor market. The results for immigrants are often smaller and not statistically significant. We show that the transitory education effects among native children are isolated to those who leave the tipped areas, suggesting that they may be driven by short-term disruptions caused by moving.

[a] This is a substantially revised version of Böhlmark and Willén (2017). We would like to thank Alexandre Mas and two anonymous referees for extremely helpful comments. We would also like to thank Matz Dahlberg, Dave Donaldson, Maria Fitzpatrick, Helena Holmlund, Michael Lovenheim, Jordan Matsudaira, Douglas Miller, Zhuan Pei, Jesse Rothstein and Alex Solis, as well as seminar participants at IFAU in Uppsala, Cornell University, SOFI, Stockholm University and the 2016 ZEW Workshop on Assimilation and Integration of Immigrants in Mannheim for valuable comments and suggestions on earlier versions of this paper. We further thank David Card, Alexandre Mas and Jesse Rothstein for sharing their program codes and Jan Amcoff for sharing the density of the Swedish SAMS areas. Alexander Willén gratefully acknowledges financial support from the Mario Einaudi Center at Cornell University and Dr. Tech. Marcus Wallenberg Foundation [2015-040]. Anders Böhlmark is grateful to the Swedish Research Council for Health, Working Life and Welfare (FORTE) [2013-064] for financial support.

1. Introduction

Ethnic and racial residential segregation are persistent features of society that generate considerable policy concerns. These concerns stem from the potential for segregation to fuel an unequal allocation of resources and opportunities across space that leads to the development of parallel societies, poses a threat to social cohesion and may negatively impact educational attainment and labor market outcomes. Despite a large theoretical literature discussing how segregation may affect individual outcomes, very little empirical work credibly addresses this question.

In this paper, we use detailed administrative data from Sweden to examine how ethnic residential segregation affects short- and long-term education and labor market outcomes of non-Western immigrants and natives. Over the past 60 years, Sweden has transformed from one of the world's most ethnically homogeneous countries to one where 22% of the population is either born abroad or has a foreign-born parent, making it an interesting case for the study of residential segregation (Statistics Sweden 2015).

The key challenge with identifying causal effects of neighborhood composition is selection: Individuals sort across neighborhoods for reasons that are unobserved by the researcher but relevant as determinants of individual outcomes. Such nonrandom selection leads to invalid inference in correlational studies since individuals in neighborhoods with different population compositions are not comparable even after adjusting for differences in observable characteristics. To overcome this issue, we borrow theoretical insight from the one-sided tipping point model used by Card, Mas and Rothstein (2008) to look at race-based tipping in the US.[1] This model predicts that residential segregation can arise due to social interactions in white preferences: once the minority share in a neighborhood passes a certain "tipping point," the neighborhood will be subject to white flight and avoidance,

[1] Card et al. (2008) derive the one-sided tipping-point model from a theory of neighborhood choice by Becker and Murphy (2000). However, several alternative models of neighborhood choice suggest similar types of behavior (see Card et al. 2011). The first formal model on the tipping phenomenon is Schelling (1971).

causing a discontinuity in white population growth.[2] Using the tipping threshold as a source of exogenous variation in ethnic population composition, we offer a novel solution to the selection issue and provide new evidence on the effect of neighborhood segregation on individual outcomes.

In a first step, we use administrative data from 1990 to 2000 to extend the work of Card et al. (2008) to look at immigrant-based tipping in Sweden's metropolitan areas - Stockholm, Gothenburg and Malmo. This exercise uses a regression discontinuity design to examine if neighborhoods on opposite sides of the tipping point in 1990 experience significant differences in native population growth between 1990 and 2000. The candidate tipping point we use is the immigrant share at which neighborhood native population growth equals the average city-specific native growth. We use this point because the one-sided tipping point model predicts neighborhoods with immigrant shares below the threshold to experience a faster-than-average native population growth and neighborhoods above the threshold to experience a relative decline.

We find robust evidence that the dynamics of segregation between immigrants and natives in Sweden's metropolitan areas is characterized by tipping. Specifically, we find that native population growth between 1990 and 2000 drops discontinuously by 9-14 percentage points among neighborhoods with immigrant shares just above 18% in 1990. Consistent with the one-sided tipping point model, our results show that neighborhoods below the threshold experience faster-than-average and stable native growth. Neighborhoods above the threshold experience a relative decline. The tipping behavior is driven exclusively by native aversion toward non-Western immigrants: the effects disappear when the model is re-estimated using Western immigrants.[3]

After having found support for immigrant-based neighborhood tipping in the metropolitan areas of Sweden, we disaggregate the data to the individual level and use the

[2] Here and throughout the paper, we follow Card et al. (2008) in defining the growth rate of the native population as the change in native population between 1990 and 2000 expressed as a fraction of total neighborhood population in 1990.
[3] This is consistent with existing literature, which suggests that segregation in Sweden is isolated to that between non-Western immigrants and the rest (Le Grand and Szulkin 2003).

tipping threshold as an instrument for changes in ethnic population composition to provide evidence on how neighborhood segregation affects individual outcomes. The intuition behind our research design is that individuals who resided in neighborhoods in 1990 that were just above the threshold should be very similar to individuals in neighborhoods that were right below the threshold on both observable and unobservable dimensions. However, individuals who lived in neighborhoods located just above the threshold will be exposed to tipping, and this will affect the ethnic population composition of the neighborhoods that they live in. It is important to note that tipping might affect other neighborhood characteristics as well, and that our estimates capture the combined effect of all these changes (we explore this further in Section 7.3). However, the scarcity of empirical work on the causal effect of residential segregation coupled with the importance of understanding the effects of tipping - which is a phenomenon that concerns many and densely populated neighborhoods - underscores the value of our study and highlights its contribution to the literature.[4]

The source of variation we exploit comes from within city across neighborhood deviations in immigrant share from the tipping point in 1990. The main assumption we invoke is that treatment assignment is as good as random around the identified threshold in 1990, so that individuals in neighborhoods just below the threshold in 1990 are comparable to individuals in neighborhoods just above the threshold in 1990. Though it is not possible to test this assumption directly, we demonstrate that there are no discontinuities in baseline characteristics of natives and immigrants at the threshold.

Our reduced form estimates identify adverse effects on the educational attainment of native children. As a percentage of the control mean, we find a 4.2% reduction in national GPA percentile ranking at age 16.[5] These effects persist as the children enter high school.

[4] In Section 5 we show that very few neighborhoods in the Swedish metropolitan areas can be categorized as fully segregated (Table 1), and in Section 6 we demonstrate that the identified tipping points are very close to the mean immigrant share across the metropolitan areas. This is thus a margin that is relevant to many communities, and it is therefore of independent policy interest to understand the consequences of segregation at this margin.

[5] The control mean is defined as the average value of the outcome variable among individuals that lived in neighborhoods in 1990 that were just to the left of the threshold.

The results for immigrants are often smaller and not statistically significant. We find no evidence that the short-term education effects carry over to the labor market.

To more fully understand the mechanisms through which these effects operate, we provide two supplemental analyses. First, we study how tipping affects the more general neighborhood environment that individuals are exposed to, since the tipping phenomenon may affect other factors than neighborhood native population growth. We find that tipping has a statistically and economically significant effect on the sociodemographic population composition of the neighborhoods, but that it does not affect the labor market opportunities and economic wellbeing of those neighborhoods. Second, we separately examine the effects on natives who "flee" the tipped neighborhoods and on natives who stay in the tipped neighborhoods. These results show that the adverse effects on the young natives education outcomes only are present among those who leave the tipped neighborhoods. Though one has to interpret these results cautiously since the decision to leave a neighborhood is endogenous, these results are consistent with the idea that the negative education effects among natives represent disruption effects caused by moving.

This is the first paper to exploit the tipping threshold as a source of exogenous variation to estimate the effect of neighborhood composition on individual outcomes. It contributes to the literature in several important ways. First, we provide a novel solution to the identification issue caused by selective sorting across neighborhoods. The application of this approach is not limited to looking at neighborhood effects and provides an interesting direction for future work on peer effects in the workplace and school. Second, while previous literature has focused on segregation of African-Americans, non-white Hispanics and refugees, this paper looks at a more heterogeneous group – non-Western immigrants (O'Flaherty 2015).[6] Given the current migration crisis in Europe, this is a group of great policy interest. Third, our identification strategy permits an investigation of

[6] Western immigrants are defined as individuals born in, or with at least one parent born in: Norway, Denmark, Finland, Iceland, Belgium, France, Ireland, Luxemburg, the Netherlands, Great Britain and Northern Ireland, Germany, Austria, Switzerland, Israel, the United States, Canada or Oceania.

segregation effects among natives, something we know very little about. Finally, while prior research has been constrained to analyzing short-term outcomes, the rich Swedish registry data enables us to follow individuals over time and investigate long-run effects.

The rest of this paper proceeds as follows: Section 2 provides a brief background on residential segregation in Sweden and relates it to that in the US, Section 3 discusses previous research on the topic, Section 4 presents our empirical strategy and Section 5 introduces the data. All results are shown in Sections 6 and 7. Section 8 concludes.

2. Background

2.1 Ethnic Residential Segregation in Sweden

During the past 60 years, Sweden has transitioned from a homogeneous to a heterogeneous society with a large immigrant base. The share of foreign-born individuals has increased from 2.8% in 1950 to 17% in 2015, and the number of non-Western foreign-born residents has increased more than twenty-fold over the same time (Appendix Table B-1). Currently, immigrants as a share of the total population in Sweden marginally exceeds that of the US, and many similarities can be drawn between the two countries. First, immigrants are spatially concentrated, and the probability of residing in an ethnic neighborhood in Sweden (0.42) is similar to that in the US (0.48) (Edin et al. 2003).[7] Second, both countries have experienced changing immigration patterns, from in-migration of Europeans to in-migration of individuals from less developed countries. As a result, immigrants have become distinctly different from natives (Chiswick and Miller 2005). Third, both countries experience disparities across ethnic groups with respect to education and labor market outcomes. In Sweden, OECD estimates suggest that the immigrant-native labor market differential is one of the largest across all member states, and recent PISA results show a 0.8 standard deviation gap in the test score distribution between natives and immigrants in math, science and reading (Åslund et al. 2011).

[7] An ethnic neighborhood is defined as a neighborhood in which the share of the neighborhood population with a specific ethnicity is at least twice as large as the share of the national population with that ethnicity. Note that the US probability is based on information from 1979, while the Swedish probability is based on data from 1997.

There are also important differences between the US and Sweden: while there are several layers of ethnic and racial segregation in the US, both across nativity status and minority groups, segregation in Sweden is restricted to that between non-Western immigrants and the rest (Le Grand and Szulkin 2003). Further, there are major differences in source countries. While Sweden has a large inflow of immigrants from the Middle East and Europe, the US has large inflows from Central America, the Caribbean and Asia.[8] Finally, the share of refugees is larger in Sweden.[9] Sweden is therefore often characterized as subject to push-migration rather than by the pull-migration present in the US.[10]

2.2 Costs and Benefits of Residential Segregation

A common finding in the literature is the existence of a correlation between a group's spatial position and socioeconomic well-being. This has motivated researchers to investigate the costs and benefits associated with residential segregation (Cutler and Glaeser 1997; Borjas 1999; Edin et al. 2003; Cutler et al. 2008). The large theoretical literature within this field point to the existence of both negative and positive mechanisms, and the resulting predictions of the effects of segregation are therefore ambiguous.

In terms of costs, existing literature suggests that ethnic residential segregation may negatively affect the desire to acquire host country specific human capital, such as language skills (Chiswick 1991; Lazear 1999). This may restrict immigrant job opportunities, in particular if the lack of such skills leads to a hesitation to explore jobs outside the neighborhood (Borjas 2000). Further, it could inhibit immigrant youth from advancing through the educational system at the same pace as natives due to inadequate proficiency in the language of instruction. Concurrently, native youth who live in neighborhoods with a high concentration of immigrants might be adversely affected if the resources at their local schools are directed toward aiding immigrants in acquiring language skills (Gould et al. 2009).

[8] In 2010, Sweden and the US did not share a single country on their top-10 source country lists (Table B-1).
[9] In 2014, 0.15 percent (491,730) of the US population was made up of refugees and asylum-seekers. In Sweden, this figure was 2.04 percent (198,342). See UNHCR (2015).
[10] See Zimmermann (1996) for a discussion of pull- and push-migration.

Residential segregation may also reduce the quality of public and private services, especially if such segregation is accompanied by an outflow of high-quality workers (Farley et al. 1994; Andersson 1998; Charles 2000). Given that the tipping phenomenon is driven mainly by native flight and avoidance, there could be sizable effects flowing through this channel, particularly if this behavior is isolated to natives of high socioeconomic status.[11]

Finally, evidence from the US suggests that neighborhoods with high ethnic concentration tend to be far removed from the suburban areas that experience job growth (Ihlanfeldt and Sjoquist 1998). According to the spatial mismatch hypothesis, the difficulty of expanding beyond neighborhood networks can cause adverse labor market effects by raising both job search and commuting costs (Kain 1968; Ihlanfeldt and Sjoquist 1998). Even though high quality transportation systems coupled with less rapid shifts in job opportunities to the suburbs make this theory less applicable to Western Europe, we are aware of no Swedish research on this hypothesis and can therefore not rule it out (Muster and Andersson 2006).

Although the majority of theories concerned with segregation predict adverse effects on immigrants, conventional social interaction models suggest that an expansion of ethnic networks may generate beneficial effects through two channels: information and norms (Bertrand et al. 2000). With respect to the former, the expansion of ethnic networks may facilitate the acquisition of vital information pertaining to education, job opportunities and social welfare programs (Patacchini and Zenou 2012; Munshi 2003). With regard to the latter, norms may improve immigrant outcomes through the transmission and sharing of work ethics and attitudes towards welfare (Borjas 1995; Glaeser, Sacerdote, and Scheinkman 1996; Bertrand et al. 2000; Åslund and Fredriksson 2009).[12]

[11] However, the direction and magnitude of the effect flowing through this channel is subject to some uncertainty, since increased segregation may also benefit and attract businesses that target immigrants.

[12] It is not clear that the effects flowing through these channels must be positive. Specifically, beneficial effects would exist only if the information (norm) benefit of expanded ethnic networks outweighs the information (norm) loss associated with a reduction in exposure to the native population.

In addition to the above mechanisms, existing research suggests that increased segregation may prolong the assimilation process, and that there thus may be treatment heterogeneity by group characteristics (Cutler et al. 2008). Specifically, if immigrants separated from majority neighborhoods revert to the native mean more slowly, then immigrants with worse labor market and education characteristics than natives may suffer while immigrants with better characteristics may benefit. Several papers have examined this hypothesis with respect to education- and skill-level, and the results are consistent with this hypothesis (Borjas 1999; Edin et al. 2003; Cutler et al. 2008).

The above discussion explains why the net effect of residential segregation is difficult to predict. By using the tipping threshold as a source of exogenous variation to examine the effect of neighborhood ethnic population composition on individual outcomes, we provide new empirical evidence on this question. However, the parameter that we estimate differs slightly from that discussed above as it captures the effect of all changes in neighborhood composition that individuals experience due to tipping, including any effects caused by moving. The scarcity of empirical work that credibly investigates the effects of residential segregation, combined with the independent policy interest in the effects of tipping, underscores the value of our empirical investigation.

3. Prior Empirical Research

Research on residential segregation falls within the literature on neighborhood effects, and the key econometric challenge associated with analyzing such effects concerns selective sorting across neighborhoods. Researchers have tried to overcome this problem using several identification strategies, ranging from randomized control trials (Katz et al. 2001; Kling et al. 2007; Chetty et al. 2015) and quasi-experiments (Jacob 2004) to propensity score matching (Harding 2003) and the use of instrumental variables (Cutler and Glaeser 1997).[13] The non-monolithic nature of neighborhood effects has led to substantial

[13] Some of the most credible estimates are from the MTO Experiment, in which families in public housing were assigned housing vouchers through a lottery, encouraging moves to areas with lower poverty rates (Sanbonmatsu et al. 2007).

heterogeneity in results across these studies, and no clear consensus has been reached on how neighborhoods affect individual outcomes (Cutler et al. 2008). Within this field of research, residential segregation has been one of the most popular subjects to examine, and this literature follows four distinct lines.

The first strand attempts to solve the endogeneity issue through aggregation to the city level (Cutler and Glaeser 1997; Collins and Margo 2000; Card and Rothstein 2007; Cutler et al. 2008; Quillian 2014). This approach is based on the assumption that neighborhood choice is endogenous to individual outcomes, but city choice is not. If correct, one can overcome the endogeneity bias by using cross-city differences in segregation as identifying variation. However, this assumption does not align with empirical evidence on migration patterns (Chiswick and Miller 2004), and several researchers have complemented this approach with additional empirical methods. For example, Cutler et al. (2008) constrain their analysis to the effect of location early in life on adult outcomes, exploit instrumental variable strategies and use fixed effects models.[14] Results from this strand are mostly negative, though some papers find mixed results (Collins and Margo 2000; Cutler et al. 2008).[15]

A second strategy limits the analysis to the effect of residential segregation early in life on adult outcomes (e.g. Cutler and Glaeser 1997; Borjas 1995; Cutler et al. 2008).[16] The assumption underlying this method is that parents choose place of residency, and if that choice is uncorrelated with unobserved characteristics that affect the children's adult outcomes, parental neighborhood choice can be used to estimate the effect of segregation among children. Although estimates using this approach suggest that immigrants are adversely affected by segregation, it is likely that parental residential choice is driven in

[14] The fixed-effects analysis uses country-of-origin and MSA fixed effects to compare outcomes between groups that are more or less segregated within a city relative to their own group-level averages. Their IV analysis uses mean years since migration for group members within a MSA as an instrument for segregation. Though informative, it is important to note that the authors do not look at the effect of segregation on natives, and they only focus on individuals between the ages of 20 and 30. Our paper addresses both of these limitations.

[15] Cutler et al. (2008) find heterogeneous effects on the skill dimension, with individuals at the bottom of the skill distribution suffering negative effects and those in the right-tail of the distribution benefitting.

[16] Borjas (1995) estimates the effect of ethnic externalities and neighborhood effects in the intergenerational transmission process and thus focuses on questions distinct from the ones that we investigate in this paper.

part by unobserved family characteristics that also affect the offspring's adult outcomes.

The third attempt to overcome the endogeneity problem has been to exploit spatial dispersal policies on refugees and asylum-seekers that generate plausibly exogenous variation in initial residential location. These policies allocate newly arrived refugees to districts based on certain observable characteristics, and if this allocation is random with respect to unobserved characteristics that also affect the outcomes, these policies can be used to estimate causal segregation effects. However, existing spatial dispersal studies have mainly focused on examining the effects of residing in an area with individuals from the same source country, as an analysis on the broader policy issue of residential segregation would require a stronger set of assumptions (Edin et al. 2003; Damm 2009; Åslund et al. 2011; Beaman 2012).[17] With the exception of Beaman (2012), these studies suggest that ethnic enclave size has a positive effect on educational and labor market outcomes.[18] Grönqvist et al. (2016) is the only paper to use these policies to examine the effect of growing up in a neighborhood with a high concentration of immigrants and finds that increased exposure leads to an increase in crime. Unfortunately, this method is restricted to looking at refugees and asylum-seekers. Further, this approach does not allow the authors to study how natives are affected by changes in immigrant concentration.

In addition to these three strands of literature, Ananat (2011) attempts to overcome the selection issue through a novel identification strategy that instruments African-American residential segregation in the 20th century using 19th century railroad configurations.[19] The results suggest that black residential segregation reduces human capital accumulation among blacks and reduces human capital inequality among whites. However, it is not clear the effects associated with ethnic residential segregation of non-Western immigrants can be inferred from these results.

[17] See Åslund et al. (2011) for a discussion.
[18] Åslund et al. (2011) and Beaman (2012) further find substantial heterogeneity in treatment effects: Åslund et al. (2011) find the positive effects to increase in the number of highly educated adults of the same ethnicity, and Beaman (2012) find that tenured co-nationals improve employment prospects and increase wages.
[19] Cities that were subdivided by railroads into a greater number of neighborhoods in the 19th century became more segregated during the great migration of the 20th century.

Our study offers a solution to the selection issue that differs from the strategies discussed above and that has never been used before. However, a number of studies have performed the first part of our estimation procedure, investigating the existence of tipping-induced segregation (Card et al. 2008; Card et al. 2011; Easterly 2009; Aldén et al. 2015; Ong 2015).[20] With the exception of Easterly (2009) that relies on a method distinct from that used by Card et al. (2008), these studies have found evidence in favor of immigrant-based tipping in Sweden and race-based tipping in the United States.[21, 22] While there is great value in examining the validity of the one-sided tipping point model, using this model as an instrument for changes in neighborhood composition represents an important contribution to the literature as it allows us to answer questions about neighborhood effects that prior research has struggled to address.

4. Empirical Methodology

The first part of our analysis extends the work of Card et al. (2008) to Sweden's three metropolitan areas. However, rather than looking at race-based tipping, we examine immigrant-based tipping. This analysis builds exclusively on the one-sided tipping point model, and a formal derivation of the empirically testable implications of this model is available in Card et al. (2008).[23] To understand our empirical method it suffices to know that the model predicts segregation to arise due to social interactions in native preference: once the immigrant share in a neighborhood exceeds a critical point, the neighborhood will experience both native flight and avoidance, causing a discontinuity in native population

[20] The tipping point literature is not isolated to looking at residential segregation. For example, Pan (2015) applies the same model to look at the dynamics of gender discrimination in the workplace.

[21] Looking at Malmo, Gothenburg and Stockholm, as well as 9 smaller cities, Aldén et al. (2015) find support for the tipping phenomenon in Sweden. However, their results cannot be compared to ours: they do not include children younger than 16 years old, do not account for second-generation immigrants and use a different definition of immigrants (individuals born outside Europe). They also estimate tipping points using a method that has a tendency to identify tipping points off of outliers (Card et al. 2008), particularly in smaller cities.

[22] Using census-tract data for US metropolitan areas from 1970 to 2000, Easterly (2009) finds that white flight is more pronounced in neighborhoods with a high initial share of whites. To the best of our knowledge, Ong (2015) is the only paper that has examined this question outside of Sweden and the US, and the author fails to find support for the tipping phenomenon in the Netherlands.

[23] The one-sided tipping point model is an alternative to the original model outlined by Schelling (1971). Schelling argues that integrated neighborhoods are inherently unstable and that social interactions in preferences will generate a completely segregated equilibrium. This can be seen as a two-sided tipping point model in which small changes in neighborhood minority composition will generate either white flight or minority flight. Card et al. (2011) compares the two models and finds that the one-sided tipping point model fits the data better. Specifically, their results show that neighborhoods with minority shares below the tipping point are relatively stable while neighborhoods above the identified tipping points are subject to significant white flight. This is consistent with our findings in Figure 1.

growth in the neighborhood. This may occur due to, for example, individuals' seeking to minimize their interactions with other-race residents (Massey and Denton 1998) or because they associate such areas with lower quality services, worse schools and higher crime rates (Krysan et al. 2008; Bayer et al. 2007). The implication of this prediction is that native population growth can be modeled as a smooth function of the immigrant share, except at the tipping point.

4.1 Identifying the Location of the Tipping Points

We follow Card et al. (2008) and assume that the tipping point is city- and decade-specific, and focus on decadal change in neighborhood population composition between 1990 and 2000.[24] To identify the location of the tipping point, we note that neighborhoods with immigrant shares below the tipping point should experience a faster-than-average native growth while neighborhoods above the threshold should experience a relative decline. One possible tipping point value is therefore the immigrant share at which neighborhood native population growth equals the average city-specific growth rate (Card et al. 2008).

To identify this point, we fit the difference between the neighborhood's decadal native growth rate and the city's mean growth rate of natives to a quartic polynomial in neighborhood base year immigrant share, measured as the fraction of non-Western first and second generation immigrants in the neighborhood.[25] As global polynomial models are sensitive to outliers, we restrict the analysis to neighborhoods with less than 60% immigrant shares:[26]

$$Dn_{sm,00} - Dn_{m,00} = f\left(i_{sm,90}\right) + \varepsilon_{sm,00}, \qquad (1)$$

where $Dn_{sm,00} = \frac{N_{sm,00} - N_{sm,90}}{P_{sm,90}}$ and denotes the change in native population N in

[24] 1990 is the first year for which we have all the data necessary for our analysis.

[25] We focus on non-Western immigrants as Western immigrants are not visible minorities and do well on the Swedish labor market (Le Grand and Szulkin 2003). Thus, it is unlikely that increases in Western immigrant shares cause native flight. We provide empirical support for this assertion in Section 6.

[26] The 60% immigrant share restriction is identical to that in Card et al. (2008) and is chosen based on visual inspection of the data to prevent outliers from affecting the identification of the tipping points. Our results are not significantly affected by changing this restriction to 50% or 70%.

neighborhood *s* and metropolitan area *m* between *1990* and *2000,* measured as a fraction of total neighborhood population P. $Dn_{m,00} = \frac{N_{m,00} - N_{m,90}}{P_{m,90}}$ and denotes the change in native population *N* in metropolitan area *m* between *1990* and *2000,* measured as a fraction of total metropolitan population *P*. *f()* is a quartic polynomial in base year neighborhood immigrant share *(i)* and $\varepsilon_{sm,00}$ is the error term. The root of this polynomial satisfies the tipping condition: $Dn_{sm,00} - Dn_{m,00} = 0$. This root is our candidate tipping point.[27] Appendix Figure B-1 illustrates how the location of the tipping point is derived based on equation (1) for a hypothetical city.

4.2 Estimating the Magnitude of the Discontinuity

To determine if there is a sufficient discontinuity in the decadal growth of neighborhood native population at the threshold to consider it a genuine tipping point, a replication of Card et al. (2008) requires that we estimate the following model:

$$Dn_{sm,00} = f\left(i_{sm,90} - i^*_{m,90}\right) + d_m \mathbf{1}\left[i_{sm,90} > i^*_{m,90}\right] + \tau_m + X_{sm,90}\beta + \varepsilon_{sm,00}, \qquad (2)$$

where *f()* is a quartic polynomial, $i_{sm,90} - i^*_{m,90}$ is the relative distance between a neighborhood's immigrant share and the identified metropolitan-common tipping point in the base year, $d_m \mathbf{1}[i_{sm,90} > i^*_{m,90}]$ is an indicator equal to one if the neighborhood had an immigrant share greater than the tipping point in the base year, X is a vector of neighborhood covariates and τ_m are metropolitan fixed-effects.[28] $d_m \mathbf{1}[i_{sm,90} > i^*_{m,90}]$ is the variable of interest. d_m captures the change in native growth between 1990 and 2000 caused by having an immigrant share greater than the tipping point in 1990.

Although equation (2) represents our preferred model specification, in Section 6 we also show results from a modified version of this equation where we instrument the

[27] In the event of several roots, we follow Card et al. (2008) and pick the one with the most negative slope. To ensure consistency with Card et al. (2008), we treat this as a two-step procedure. After we identify a candidate tipping point (CTP), we repeat the procedure using only neighborhoods with $abs\left(i_{sm,90} - CTP\right) < 10$ to zero-in on the true tipping point.

[28] Covariates are not necessary in a regression discontinuity framework. However, they can reduce the sampling variability and improve precision (Lee and Lemieux 2010). In results not shown we have estimated our models without including our control variables. Consistent with Lee and Lemieux (2010), this does not affect our estimates.

fraction of the decade that the neighborhood was exposed to tipping with whether the neighborhood was above or below the candidate tipping point in the base year. The rationale underlying this alternative model is that we analyze decadal change in neighborhood native population based on the neighborhood's distance to the tipping point in 1990. A regular inflow of immigrants to control neighborhoods may cause control neighborhoods close to the tipping-point to move beyond the threshold – and tip - later in the decade. With respect to the current study, we find that 62 neighborhoods (12%) move from below the tipping point in 1990 to above the tipping point between 1991 and 2000 (Appendix Table B-2). We also find that 8 neighborhoods (1.5%) just above the threshold in 1990 move below the threshold at some point between 1991 and 2000. These neighborhoods will attenuate our point estimates. By instrumenting the fraction of the decade that the neighborhood was exposed to tipping with whether the neighborhood was above or below the candidate tipping point in the base year, we can get a sense of the attenuation caused by crossovers while maintaining the quasi-experimental identification that the tipping model provides.[29]

We follow Card and Lee (2008) and cluster the standard errors on distinct values of the running variable. To evaluate the robustness of our results, we estimate several modified versions of equation (2). First, although the global polynomial RD approach offers greater precision than the nonparametric RD approach, it is difficult to identify the correct functional form. We therefore examine the sensitivity of our results to alternative polynomial specifications, and to the use of local linear regressions. Second, we acknowledge that the location of the tipping point may be subject to measurement error which makes it harder to detect an effect, and we therefore complement our baseline analysis with donut-style regression discontinuity models that allow tipping to occur within a small range around the threshold rather than exactly at that point.

[29] While this exercise is useful for understanding how crossovers may attenuate our results, we prefer equation (2). The reason is that it is not clear, in particular when looking at individual outcomes, that the fraction of the decade that the neighborhood was exposed to tipping is the correct treatment.

A random 2/3 of neighborhoods within each metropolitan area is used for the data-intensive process of identifying the location of the tipping points via equation (1). To estimate the magnitude of the discontinuities and determine if the identified thresholds represent genuine tipping points, we rely on the 1/3 of neighborhoods within each metropolitan area not used to identify the location of the threshold.[30] This split-sample procedure is used due to specification search bias – the magnitude of the discontinuity will have a non-standard distribution under the null hypothesis of no structural break if the same sample is used to identify the tipping point and estimate the discontinuity (Card et al. 2008; Leamer 1978). As a consequence, conventional test statistics will reject the null hypothesis of no discontinuity too often. Using two random subsamples means that the samples are independent and will have a standard distribution even under the stated null hypothesis (Card et al. 2008).

4.3 Individual Outcomes

After having found support for immigrant-based tipping in Sweden's metropolitan areas, we disaggregate the data to the individual level and estimate the effect of neighborhood ethnic population composition on key outcomes, using the tipping threshold as a source of exogenous variation. We perform this analysis separately for non-Western immigrants and natives from three different age groups: those born 1980-1990, those starting school between 1980 and 1990, and those who have completed their education between 1980 and 1990 (born 1948-1958).[31] We estimate the following reduced-form model:

$$Y_{rsm,t} = f\left(i_{sm,90} - i^*_{m,90}\right) + d_m \mathbf{1}\left[i_{sm,90} > i^*_{m,90}\right] + \tau_m + \partial_r + X_{rsm,90}\beta + \varepsilon_{rsm,t}, \quad (3)$$

where $Y_{rsm,t}$ is an outcome at time t for resident r that lived in neighborhood s in

[30] We restrict attention to the three largest cities in part because they are the only metropolitan areas in Sweden, and in part due to power concerns. The ten largest cities excluded from our sample have an average of less than 70 neighborhoods. We would therefore have less than 50 neighborhoods to identify thresholds from, and less than 24 neighborhoods to use for identifying the magnitude of the discontinuity in these areas.

[31] Our decision to perform cohort-specific analyses is guided by Chetty et al. (2015), who show that there may be substantial birth cohort heterogeneity with respect to neighborhood effects.

metropolitan area m in 1990, X is a vector of individual-level covariates measured in 1990 and ∂ are birth year fixed effects. To avoid specification search bias, we restrict our sample to individuals in the 1/3 of neighborhoods within each metropolitan area not used to identify the location of the threshold when.

The intuition behind our research design is that individuals that lived in neighborhoods in 1990 that were just above the threshold should be very similar to individuals that lived in neighborhoods in 1990 that were right below the threshold. However, individuals in neighborhoods above the threshold will be subject to tipping, and this will affect the neighborhood population composition that they are exposed to. Thus, we compare individuals who resided in comparable neighborhoods in 1990 but experienced significantly different neighborhood population compositions in the following decade due to very small initial differences in neighborhood immigrant shares. It is important to note that tipping might affect other neighborhood characteristics as well, and that our estimates capture the combined effect of all changes in neighborhood composition that individuals experience due to residing in a neighborhood in 1990 that was above the tipping point in that year.

The source of variation we exploit comes from within city across neighborhood deviations in immigrant share from the tipping point in 1990. The main assumption we invoke is that treatment assignment is as good as random around the identified threshold in 1990, so that individuals that lived in neighborhoods in 1990 that were just below the threshold are comparable to individuals that lived in neighborhoods in 1990 that were just above the threshold. Though it is not possible to test this assumption directly, we demonstrate that there are no discontinuities in baseline characteristics of natives and immigrants at the threshold.

In addition to our main assumption, the validity of our estimation strategy requires that the tipping points are correctly estimated, that there are no coincidental shocks affecting neighborhoods once they hit the tipping point that also affect the outcomes of

interest, and that the functional form used to model the relationship between the conditional mean of the outcome and running variable is correctly specified. In Section 7, we report results from several robustness checks and diagnostic tests that show that our data are consistent with these assumptions.

5. Data

We rely on detailed population-wide administrative data drawn from four registries of the IFAU database, originally collected by Statistics Sweden. The first registry is *The Longitudinal Database for Education and Labor Market Outcomes (LOUISE)*, which contains annual socioeconomic and demographic information on all residents between the ages of 16 and 65. The data also contain information on the number of children below the age of 16 in every household, allowing us to incorporate children into the estimation of tipping points. To examine if the tipping points are associated with discontinuities in neighborhood population composition we use data from 1990 to 2000, focusing on decadal change in native, immigrant and total population growth. To evaluate if tipping impacts individual outcomes we append LOUISE data from 2001 to 2011. This allows us to investigate long-run effects of the tipping phenomenon on individual outcomes.

The ability to follow individuals over time and across space is crucial to our analysis, as our estimation strategy requires knowledge of each individual's residential history as well as his/her labor market and education progression. The demanding data requirement of this strategy is a main reason why this analysis has not been performed in the past. Another crucial data component is the neighborhood classification system (*SAMS)*. *SAMS* is the most detailed geographic division in Sweden and divides the country into municipality-confined blocks, with a mean size of 1,000 individuals. This represents a finer level of geographic division than that used by Card et al. (2008).

The LOUISE registry contains information on education, labor market participation, income and welfare program participation; all of which we use as outcome variables when estimating the reduced-form effect of neighborhood composition via equation (3). We

supplement these measures with data from the *Grade 9 Registry* and the *High School Registry*, which provide information on the academic performance of all individuals at the compulsory and high school levels.

The fourth registry we use is *The 2009 Multigenerational Registry (FLERGEN)*, which links all individuals that resided in Sweden at some point after 1961 to their family members. We use these data for two purposes. First, by linking Swedish-born individuals in LOUISE to their parents via FLERGEN, we are able to account for second-generation immigrants. Second, by linking individuals from the 1975-1990 cohorts to their parents via FLERGEN, we can use the parental characteristics of these individuals when they were children to identify where they lived before they turned 16.

Consistent with prior literature, we impose three sample restrictions. First, we exclude neighborhoods with growth rates five standard deviations above that of the city, as there may be coincidental secular trends that bias the estimates in these areas (Card et al. 2008). Second, we exclude neighborhoods with less than 200 residents, as a small change in the number of immigrants can cause tipping in these areas (Aldén et al. 2015). Finally, we drop neighborhoods that only exist for part of the decade. These restrictions reduce our sample by 886 neighborhoods, 762 of which have less than 200 residents.[32] Our final data set consists of 1,560 neighborhoods in the 3 metropolitan areas of Sweden. About 85% of the populations in these cities are included in our sample.[33]

Table 1 provides a breakdown of total and native growth rates stratified by baseline immigrant share. Most neighborhoods have between 5 and 40 percent immigrants in 1990. That the percent of neighborhoods with more than 40% immigrants doubles between 1990 and 2000 provides suggestive evidence in favor of an increase in ethnic residential segregation over the time period under consideration in this paper.

[32] Card et al. (2008) also drop neighborhoods with ten-year white growth rates in excess of 500 percent of the total base-year population. This restriction does not lead us to drop any neighborhoods.

[33] With the exception of the growth rate variables, the excluded neighborhoods are not statistically significantly different from those included in the sample (Appendix Table B-13). The difference in growth rates between the included and excluded neighborhoods is likely driven by the relatively small size of the excluded neighborhoods, as the addition of one more individual can have a substantial effect on those measures. Including these neighborhoods would greatly increase the risk of identifying tipping points based on outliers.

Table 2 provides summary statistics of the individuals used for the second part of our analysis, stratified by birth cohort. Columns 1-3 provide descriptive statistics of all individuals in Sweden, columns 4-6 provide the same statistics for all individuals in Sweden's three metropolitan areas, and columns 7-9 show the characteristics of the individuals in these metropolitan areas after our sample restrictions have been imposed. The characteristics of the individuals included in our sample closely mirror the average characteristics of the three cities, and there are no statistically significant differences between the two groups. Thus, our final sample is strongly representative of the population in the three metropolitan areas of Sweden.[34]

6. Tipping Point Results

6.1 Baseline Estimates

Using a random 2/3 of neighborhoods within each metropolitan area, estimation of equation (1) identifies the unweighted mean tipping point across the three metropolitan areas to be located at an immigrant share of 17.94, with a standard deviation of 0.92. The tipping point is located at a slightly higher immigrant share in Gothenburg (18.99) than in Malmo (17.28) and Stockholm (17.55).[35] The mean tipping point is close to the mean immigrant share across the meteropolitan areas in 1990 (19.03). This phenomenon thus occurs at an immigrant share relevant to a large number of neighborhoods.[36]

By pooling the neighborhoods and normalizing the city-specific tipping points to zero, Figure 1 (a) replicates Figure V of Card et al. (2008) for Sweden's meteropolitan areas and provides preliminary evidence of a discontinuity in neighborhood native population growth at the threshold. Only the 1/3 of neighborhoods within each metropolitan area not used for estimating the location of the tipping points are used for this depiction. The dots show mean change in native population between 1990 and 2000,

[34] A concern with the data is that it fails to account for undocumented immigrants. However, recent estimates suggest the upper bound of undocumented immigrants in Sweden to be 35,000 during the time period that we analyze (SOU 2011). This represents 0.37% of the total population (or 1.02% of the immigrant population).

[35] We have also estimated tipping points using the structural break method described in Card et al. (2008). See the Structural Break Appendix (page 178) for a description of this method. Although this method is very susceptible to outliers in smaller cities, it produces results consistent with our main findings: the mean tipping point is 19.9, with a standard deviation of 1.96.

[36] This assertion is contingent on the shape of the density of the fraction non-Western immigrants in the base year. For example, if this density is bimodal, this would not follow. However, Figure 2 shows that this is not the case.

grouping neighborhoods into 2% bins by the deviation in immigrant share from the tipping point in 1990. The horizontal line is the unconditional mean, and the vertical line represents the normalized tipping point. The solid line is a local linear regression fit separately on either side of the threshold weighted by the size of the neighborhoods, using an Epanechnikov kernel and a bandwidth of 4. Appendix Figure B-3 shows that this result is robust to changing bandwidth and polynomial order.

Figure 1 (a) provides strong evidence of a negative discontinuity in native growth at the tipping point. The discontinuity at the tipping point is approximately 10 percentage points. The positive and flat slope of the local linear regression fit to the left of the tipping point coupled with the downward trend to the right of the threshold is consistent with Figure V of Card et al. (2008). To explore the outliers located close to the threshold in Figure 1 (a), Figure 1 (b) replicates Figure 1 (a) but weights the size of each marker by the total base year size of the neighborhoods used to obtain that point. Figure 1 (b) shows that the noisiest dots are generated by the smallest neighborhoods in the cities, increasing our confidence in the empirical strategy.

Figures 1 (a) and (b) suggest that the dynamics of ethnic segregation in Sweden's metropolitan areas is characterized by tipping behavior. However, these figures do not allow for formal hypothesis tests. To this end, Table 3 shows estimates of equation (2) for the growth rate of natives, immigrants and natives and immigrants combined.[37] The coefficient on *Beyond TP* captures the effect of being above the tipping point in 1990 on the change in the respective outcomes between 1990 and 2000. Columns 1 through 3 show results obtained from ordinary least squares regressions while Columns 4 through 6 show results obtained from two-stage least-squares regressions, in which the fraction of the decade that the neighborhood was exposed to tipping has been instrumented with whether it was above or below the threshold in 1990.

[37] If total population size has an independent effect on individual outcomes, the negative discontinuity in total population growth shown in Table 3 may drive some of our results in Section 7. Though we are aware of no studies that examine this question, if present, we consider this part of the treatment.

The effect of tipping on the change in neighborhood native population is shown in Columns 1 and 4. The point estimates shown in these columns provide clear evidence of the existence of a large negative discontinuity, with estimated effect sizes of 9 and 14 percentage points respectively. These results are consistent with the visual evidence presented in Figure 1 (a). The statistically significantly larger effect in Column 4 is indicative of crossovers causing a slight attenuation of the point estimate presented in Column 1. Consistent with the one-sided tipping point model and the results in Card et al. (2008), Table 3 also shows that tipping had a statistically significant impact on the change in total population but not on the change in immigrant population.

6.2 Robustness Checks and Diagnostic Tests

In order to exploit the identified discontinuities to estimate the reduced-form effect of tipping on individual outcomes, we need to invoke a number of assumptions. The main assumption is that neighborhood immigrant shares move smoothly through the tipping points in the base year. If this is not the case, there may be manipulation of the running variable that could threaten the validity of our estimation strategy. In this context, such manipulation is highly unlikely as it would require coordinated action of multiple individuals or explicit government policies that keep the immigrant share right below (or above) the threshold that we have identified.

Figure 2 plots the frequency of observations by 2 percent bins in the deviation in immigrant share from the estimated tipping point in the base year. The vertical line depicts the normalized tipping points. The solid line is a local linear regression fit separately on either side of the tipping point, using an Epanechnikov kernel and a bandwidth of 4. The figure shows that there is no discontinuity in the density at the cutoff.[38] This result is robust to changes in bandwidth and polynomial order. We have also performed the McCrary (2008) density test, which fails to rejects the null that the discontinuity is zero.

[38] Although there is more mass to the left of the threshold, it is a potential discontinuity at the tipping point value, not the fact that there is more mass on one side of it, that would indicate manipulation (Lee and Lemieux 2010).

The absence of a discontinuity in the density of immigrant shares around the threshold in 1990 is encouraging but does not exclude the possibility that there is systematic sorting of individuals across neighborhoods around the threshold that affects the outcomes. We therefore estimate equation (2) for several native and immigrant characteristics measured in the base year. These covariates are determined prior to treatment and should not be subject to discontinuities at the threshold. Table 4 shows that only one of the 22 coefficients (immigrant social welfare participation) is statistically significant, and the size of this coefficient is small. We interpret these results as demonstrating that native and immigrant characteristics are smooth across the tipping point, and as evidence against treatment manipulation.

A concern specific to this analysis relates to variation in the tipping points over time. Individuals are assigned to treatment based on tipping points in 1990, and substantial fluctuation in tipping point values over time may dilute our estimates. To investigate this concern, we re-estimate equation (1) for the decade succeeding our analytical period.[39] The mean tipping point for this decade is 20.77, slightly higher than that identified for the 90-00 period.[40] This increase is driven by Gothenburg: in Stockholm and Malmo the tipping points increase by less than 0.8. We therefore re-estimate equation (2) using only neighborhoods from Stockholm and Malmo. The point estimate becomes slightly larger when Gothenburg is excluded. However, the difference between this coefficient estimate and our baseline estimate is not statistically significant, and we continue to use the full sample throughout.[41]

Another worry pertains to measurement error in the location of the tipping points, as this can smooth away true discontinuities and attenuate our results. One way to investigate if this constitutes a problem is by allowing tipping to occur within a certain range of the identified tipping point, and estimate donut-style regression discontinuity models in which

[39] Unfortunately, we lack data to perform a similar calculation for the decade preceding our analytical period.
[40] The slight increase is consistent with Card et al. (2008) and Aldén et al. (2015), as well as with Oliver and Wong (2003) who argue that exposure to immigrants in integrated neighborhoods may counter stereotypes.
[41] When Gothenburg is excluded the discontinuity in native population change becomes -0.104, with a standard error of 0.049.

neighborhoods with baseline immigrant shares within this range are excluded. Appendix Table B-3 shows the results from this exercise for five different donut-hole sizes: 0.10, 0.30, 0.50, 1.00 and 2.00. None of these specifications produce estimates that are statistically significantly different from our baseline results; measurement error in the location of the threshold is unlikely to bias our results.

An additional concern relates to our decision to define Western immigrants as natives. If natives are averse to having Western immigrant neighbors, this grouping will cause attenuation bias. To examine this possibility, we calculate new tipping points using equation (1), and then re-estimate equation (2) with immigrants defined as first-or second-generation Western immigrants (Appendix Table B-4). There is no evidence in favor of discontinuities at these alternative tipping points. This is consistent with existing literature that has found residential segregation in Sweden to be isolated to that between non-Western immigrants and the rest (Le Grand and Szulkin 2003).

A final worry relates to the functional form used to model the relationship between the conditional mean of the outcome and running variable. An incorrect functional form will cause the resulting estimator to be biased, and it thus is appropriate to explore the robustness of the results with respect to alternative polynomial orders and specifications. Appendix Table B-5 shows the results from this exercise. The point estimate is insensitive to polynomial order, the inclusion of additional controls, controlling for population density and fully interacting the polynomial with the dummy variable for whether the neighborhood was above or below the tipping point in 1990.

In addition to examining the sensitivity of our results to different model specifications, we also examine the robustness of our results to estimating a local linear regression in a given window around the threshold. The advantage of this approach is that it does not rely on observations far from the threshold for identification.[42] Both the global polynomial approach and the local linear regression approach estimate the same statistic.

[42] If the underlying function is not exactly linear in the area being examined, the local linear regression results may be associated with a substantial bias. However, given the visual evidence in Figure 1 this type of bias is unlikely.

Robustness of our results across these two specifications therefore indicates that the point estimates are not driven by specific functional forms. Appendix Table B-6 displays results from the nonparametric approach and shows that our estimates are robust to the use of local linear regressions.[43]

7. Individual Outcomes

7.1 Education Effects

We investigate three different sets of educational outcomes. First, academic performance in the 9th and final year of compulsory school (GPA and grades in the core subjects Swedish, English and mathematics). We look at subject-specific grades as the discussion in Section 2 suggests a clear link between residential segregation and the motivation and ability to acquire host country language skills. It is thus possible that student performance in Swedish is more affected than that in other languages and subjects. Second, high school performance (GPA, probability of attending a science track and the probability of enrolling in an academic (university-preparatory program).[44] Third, post-secondary educational attainment measured in 2011 (years of schooling and the probability of having attended university). We have converted the grades into yearly national percentile rankings. These outcomes span the entirety of our cohorts' educational experiences.

Baseline estimates of the effect of neighborhood population composition on educational attainment, stratified by nativity status and cohort, are shown in Table 5. Each cell comes from a separate estimation of equation (3). The point estimates displayed in the table represents the reduced-form effect of neighborhood composition on the outcome listed at the top of the column, using the tipping threshold as the source of exogenous variation. The table further shows the effect as a percentage of the control mean. It is important to note that our results in Table 3 suggest that immigrants do not respond to tipping, and that the neighborhood composition estimates for our immigrants subsample

[43] To perform this analysis, we rely on the cross-validation method proposed by Ludwig and Miller (2005) to obtain an optimal bandwidth. We have also used a bandwidth twice as large and half that recommended by the cross-validation method. The point estimates are not statistically significantly different.

[44] We have also examined these outcomes by restricting the sample to individuals who graduate on time, as segregation could affect one's decisions of when to enroll. However, this has no effect on our estimates.

are cleaner than those for natives (since one does not have to worry about a potential moving effect).

The top panel displays the reduced-form effect on the young immigrants' educational outcomes. These estimates are small and not statistically significant, suggesting that the change in neighborhood composition is not associated with negative education effects for young immigrants. There is some suggestive evidence of adverse education effects in compulsory school among immigrants in the middle cohort, though these results are only marginally statistically significant.

With respect to natives, our estimates point to adverse effects on educational attainment. As a percentage of the control mean, tipping leads to a 4.2 percent reduction in 9th grade national GPA percentile ranking (driven primarily by a reduction in the Swedish component grade), and a 2.6 percent reduction in high school national GPA percentile ranking, among natives from the young cohort.[45] There is clear evidence of adverse education effects among natives from the middle cohort as well, though these effects are only statistically significant at the compulsory school level: As a percentage of the control mean, the tipping phenomenon leads to a 2.2 percent reduction in 9th grade national GPA percentile ranking, driven by reductions in both the English and the Swedish component grades. It should be noted that the point estimates for immigrants are based on a much smaller sample, and that these estimates therefore are less precise.

To further examine the effect on cognitive and non-cognitive skills, we supplement our data with information on military tests that men took when conscription was mandatory.[46] The test scores range from 1 to 9 and were used to place individuals into military branches. Results from estimating equation (3) using these test scores as dependent variables are shown in Appendix Table B-8. The results suggest that the

[45] To understand the size of these coefficients, it is useful to place them in relation to the effects of traditional education interventions, such as class-size reductions. In a recent study, Fredriksson et al. (2013) find that a one-pupil reduction in class-size in grades 4-6 in Sweden improves 9th grade academic achievement by 0.023 of a standard deviation. Our reduced form estimates for the young native cohort therefore suggest that the adverse education effect of tipping is equivalent to that of increasing the average class size with 2 pupils.

[46] This is similar to the AFQT in the United States. See Mårdberg and Carlstedt (1993) for a description.

cognitive and non-cognitive ability of these individuals are unaffected by the change in neighborhood composition. The adverse effects identified in Table 5 are thus driven by skills or behaviors not captured by these tests.

7.2 Labor Market Effects

We focus on three labor market outcomes and estimate both intensive and extensive margin effects: *Employment income* (annual earnings from employment, excluding self-employment income but including work-related compensation from the Social Insurance Agency), *Self-employment income*, and *Government-funded benefits* (compensation from 32 social security programs, including educational grants, grants to immigrants for learning Swedish, unemployment benefits, early-retirement supplemental compensation, compensation for start-ups and compensation for voluntary military service).[47] The outcomes are measured in 2011. We transform the income variables to their natural logarithms when analyzing intensive margin effects and we convert them to dichotomous variables when examining extensive margin effects.

Panel A (immigrants) and Panel B (natives) of Table 6 display the results obtained from estimation of equation (3) for each of the labor market outcomes listed above. Looking across Panel A, there is no evidence of an effect on immigrant earnings, either on the intensive or the extensive margin.[48] We also do not find any evidence of a significant effect on the probability of being self-employed, the amount of self-employment income that an individual receives, the probability of receiving government-funded benefits, or the amount of government-funded benefits that an individual receives. These findings hold true for all three cohorts that we look at. With respect to our native subsample (Panel B), the results tell a similar story. Specifically, there is no evidence that the change in neighborhood population composition has an impact on earnings, self-employment income or government-funded benefits, either on the intensive or the extensive margin.

[47] In results not shown, we have also examined the effect on *Socialbidrag* – government assistance to individuals who earn less than the amount considered necessary for supporting oneself financially. We find no effects.
[48] Excluding work-related compensation from employment income does not alter the interpretation of the results.

The young and middle cohorts are between 21 and 38 years old when the labor market outcomes are measured, and the majority of these individuals are not on a part of their earnings profiles where yearly earnings are informative about lifetime earnings (Haider and Solon 2006; Böhlmark and Lindquist 2006). In results not shown, we have estimated equation (3) only using individuals who are between 33 and 38 when the outcomes are measured. The results are not statistically or economically significantly different from the point estimates depicted in Table 6.

7.3 Potential Mechanisms

The results in Table 5 show that the negative education effects are predominantly isolated to the young natives' short-term educational outcomes, and the estimates in Table 6 demonstrate that these adverse education effects do not carry over to the labor market.

To more fully understand the channels through which these effects may operate, we perform two supplemental analyses. First, we examine the relationship between tipping and the more general neighborhood environment, since the tipping phenomenon may affect other factors than the neighborhood native population growth rate. To this end, we reestimate equation (2) using two indices meant to capture the sociodemographic population composition and economic wellbeing of the neighborhood in 2000 as dependent variables. The *Economic Activity Index* (EA) consists of three labor market variables intended to capture the economic wellbeing of the neighborhood: average employment income, average years of education, and fraction employed. The *Sociodemographic Index* (SD) consists of four sociodemographic variables intended to capture the population composition of the neighborhood: gender balance, age profile, fraction immigrants and fraction on social welfare. For each of these indices, we use unity-based normalization to bring the values of each of the variables into the range [0,1], take their sum, and standardize this sum to have a mean of zero and a standard deviation of one. We interpret a reduction in the EA Index as a decline in the labor market opportunities and economic wellbeing of the neighborhood, while a reduction in the SD Index is indicative

of a change in the sociodemographic make-up of the neighborhood, such that it's population is becoming older, less gender balanced, have a larger fraction of immigrants and have more individuals on social welfare.

Figure 3 (a) demonstrates that there is no sign of a discontinuity in the EA Index at the threshold while Figure 3 (b) provides evidence of a discontinuity in the SD Index.[49] The discontinuity at the tipping point in Figure 3 (b) is approximately 30 percent of a standard deviation. Estimation of equation (2) shows that this discontinuity is significant at the 1 percent level (Appendix Table B-14).[50] This set of results imply that any potential effect on individual outcomes likely operate through changes in the sociodemographic make-up of the neighborhood that the individual is exposed to.[51]

The second supplemental analysis that we perform is to reestimate the short-run educational attainment results separately for those who remained in their original neighborhood throughout the decade and for those who did not. This analysis is motived by the transitory nature of the education effects, suggesting that the effects could be driven entirely by short-term disruptions caused by moving among natives. While there are clear issues with this approach (since the decision to move is endogenous), we believe that these results is informative for understanding which group of individuals is driving the results.

---

[49] Since the tipping phenomenon is partly driven by natives moving away from tipped neighborhoods, examining discontinuities in the SD and EA indices at the neighborhood level is limiting, as these are not necessarily the neighborhoods that the individuals are exposed to ten years later (due to moving). To this end, we disaggregate the data to the individual level and examine whether living in a neighborhood in 1990 that was just above the tipping point affects the neighborhood environment that the individuals are exposed to in 2000. As shown in Appendix Table B-14, the results from these individual-level regressions mirror the neighborhood results: individuals that lived in neighborhoods in 1990 that tipped that year were exposed to neighborhood environments in 2000 that had a different sociodemographic make-up, but that were otherwise similar.

[50] To further explore if immigrant-based tipping contributes to socioeconomic segregation, we stratify the native sample by income (top and bottom quartile), education (more or less than a high school diploma) and gender, and estimate equation (3) for two outcomes: (a) the probability of moving from a treatment to a control neighborhood and (b) the probability of moving from a control to a treatment neighborhood. The results from this exercise are shown in Appendix Table B-7. The point estimates show that individuals in the left-tail of the socioeconomic distribution are more likely to move from a control neighborhood to a treated neighborhood during the analysis period compared to individuals in the right-tail of the socioeconomic distribution. However, these differences are very small and often not statistically significant. This is consistent with the results in Figure 3 and Appendix Table B-14, showing a discontinuity in the SD Index, but not in the EA Index, at the threshold.

[51] Appendix Table B-9 provides statistics on the fraction of individuals who remain in their 1990 neighborhood over time. Approximately half of all individuals that lived in neighborhoods that did not tip in 1990 remain in those neighborhoods ten years later, while only 38 percent of individuals that lived in neighborhoods that tipped in 1990 remained in those neighborhoods ten years later. The lowest compliance rate can be found among natives that lived in neighborhoods that tipped in 1990, and this is expected given the results in Section 6. Further, while the difference in compliance rate between natives in tipped and non-tipped neighborhoods is relatively large (13 percentage points), the difference in compliance rates between immigrants in tipped and non-tipped neighborhoods is not (4 percentage points). This is also in line with the results in Section 6.

The results from this exercise are shown in Table 7. Looking across the table, the negative effects are entirely driven by individuals who resided in neighborhoods in 1990 that were above the tipping point but that then decided to move away from those neighborhoods at some point during the following decade. These results support the idea that the identified effects are driven entirely by short-term disruptions caused by moving. However, it is important to note that our empirical strategy does not permit us to conclude this with certainty. The reason is that we cannot exclude the possibility that those who decide to move from tipped neighborhoods are the ones that suffer the most from the compositional changes caused by tipping, and that they would have suffered similarly negative effects had they decided to remain in their original neighborhoods.

7.4 Robustness and Sensitivity Analyses

We perform a series of sensitivity checks to investigate the robustness of our results to minor alterations of the empirical model. Each of the alterations deals with a specific concern associated with our estimation strategy. The first concern relates to variation in neighborhood population density both within and across the metropolitan areas. This variation in population density will affect the level of individual exposure to immigrants and natives, and may therefore impact the effect of tipping on individual outcomes.

Second, the effects may differ for individuals in neighborhoods that are surrounded by other neighborhoods that have tipped, as individuals who live in tipped neighborhoods may work and study in neighboring areas. A specific concern is therefore that our results are driven by individuals that lived in tipped neighborhoods in 1990 that were fully surrounded by other tipped neighborhoods. To investigate this, we use Statistics Sweden's geographic neighborhood atlas to identify neighborhoods that surround areas that have tipped (Appendix Table B-11).[52] We look at whether our results are robust to the exclusion of individuals that resided in neighborhoods with 100% tipped neighbors.[53]

[52] www.scb.se/sv/Vara-tjanster/Regionala-statistikprodukter/Marknadsprofiler/Postnummer-och-SAMS-atlasen/

[53] In analyses not shown, we have stratified the sample based on the percent of neighboring areas that tipped. We do not find a difference in effect sizes.

Third, individuals in neighborhoods in the right-tail of the non-Western immigrant share distribution in 1990 may be different from individuals in neighborhoods at the margin of tipping on dimensions that equation (3) cannot control for.

Appendix Table B-15 (immigrants) and Appendix Table B-16 (natives) display results obtained from running each of the modified regressions for each of the cohorts. The first row of each panel controls for neighborhood population density, the second row shows the results obtained when we exclude individuals that resided in neighborhoods with 100% tipped neighbors in the base year and the third row displays the results when outliers are omitted. The results show that our baseline estimates are robust to these alternative model specifications.[54]

8. Discussion and Conclusion

Identifying the effects of residential segregation through empirical analysis is difficult due to selective sorting across neighborhoods, and prior research in this area has been hampered by a lack of exogenous variation in neighborhood choice. We overcome this problem by utilizing a novel identification strategy that borrows theoretical insight from the one-sided tipping point model of Card et al. (2008), using the tipping threshold as a source of exogenous variation in ethnic population composition.

We find robust evidence that the dynamics of segregation between immigrants and natives in Sweden's metropolitan areas is characterized by tipping. Specifically, we find that native population growth between 1990 and 2000 drops discontinuously by 9-14 percentage points among neighborhoods with immigrant shares just above 18% in 1990. Consistent with the one-sided tipping point model, our results show that neighborhoods below the threshold experience faster-than-average and stable native growth. Neighborhoods above the threshold experience a relative decline.

[54] The discussion in Section 2 suggests that there may also be treatment heterogeneity on the time dimension. To explore this possibility, we extend our analysis in two ways. First, we complement our analysis of the oldest cohort by looking at their labor market outcomes in 2000. Second, we look at the effect of tipping on employment income for the oldest cohort for each year between 1990 and 2011. Results from the first exercise are shown in Appendix Table B-12. Results from the second exercise are shown in Appendix Figure B-2. These results are inconsistent with the existence of heterogeneous treatment effects on the time dimension.

In the second part of our analysis, we use the tipping threshold as a source of exogenous variation to examine the effect of neighborhood composition on individual outcomes. Our results identify modest adverse education effects on native children. As a percentage of the control mean, we find a 4.2% reduction in national GPA percentile ranking at age 16. These effects persist as the children enter high school. We find no evidence that the education effects carry over to the labor market.

To better understand the mechanisms underlying these effects, we perform two supplemental analyses. First, we show that tipping affects the sociodemographic population composition of the neighborhoods that individuals are exposed to, but not the economic wellbeing of those neighborhoods. Second, we show that the short-term education effects among young natives are driven entirely by those who leave the tipped areas, suggesting that these effects may be driven entirely by short-term disruptions caused by moving. However, one has to be careful when interpreting these results as the decision to move is endogenous, and our empirical strategy does therefore not allow us to conclude this with certainty.

In terms of policy implications, our results demonstrate that social interactions in native preferences represent an obstacle to neighborhood integration. Conventional place- and people-based policy solutions to residential segregation would only have a minimal impact on reducing the prevalence of this phenomenon, and policymakers may need to look at alternative approaches that target the root cause of the problem. Our results further suggest that the effects of neighborhood population composition on individual outcomes are minimal at most. However, it is important to highlight that we study the effects of neighborhood composition on individual outcomes in a country with a very generous social policy system, and that Sweden's social policy system may mute some of the effects associated with tipping. For example, Sweden's financial equalization schemes and generous welfare policies may hedge against the anticipated quality reductions in services and institutions discussed in Section 2. Therefore, one should be careful to extrapolate

these results to other countries and settings, as variation in social policies and public institutions likely affect the results. For example, market-driven housing, property tax funded schools and a large share of unauthorized immigrants that cannot access welfare services make it possible that the effects would be different in a country such as the US.

References

Aldén, L., M. Hammarstedt, and E. Neuman (2014) "Ethnic Segregation, Tipping Behavior and Native Residential Mobility" *International Migration Review* 49(1): pp. 36-69

Ananat, E. (2011) "The Wrong Side(s) of the Tracks: The Causal Effect of Racial Segregation on Urban Poverty and Inequality" *AEJ: Applied Economics* 3: pp. 34-66

Andersson, R. (1998) "Socio-spatial dynamics: Ethnic divisions of mobility and housing in post-Palme Sweden" *Urban Studies* 35: pp. 397-428

Bayer, P., F. Ferreira and R. McMillan (2007) "A unified framework for measuring preferences for schools and neighborhoods" *Journal of Political Economy* 115(4): pp. 588-638

Beaman, L.A. (2012) "Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S." *Review of Economic* Studies 79(1): pp. 128–161

Becker, G., and K.M. Murphy (2000) *Social Economics: Market Behavior in a Social Environment*, Cambridge: Harvard University Press

Bertrand, M., E. Luttmer, and S. Mullainathan (2000) "Network Effects and Welfare Cultures" *The Quarterly Journal of Economics* 115(3): pp. 1019–1055

Borjas, G.J. (1995) "Ethnicity, Neighborhoods, and Human Capital Externalities" *The American Economic Review* 85(3): pp. 365–390

Borjas, G. J. (1999) "Immigration and Welfare Magnets" *Journal of Labor Economics* (17): pp. 607–637

Borjas, G.J. (2000) "Ethnic Enclaves and Assimilation" *Swedish Economic Policy Review* Vol. 7(2): pp. 89–122

Boustan, L. (2011) "Racial Residential Segregation in American Cities" in *Handbook of Urban Economics and Planning*, eds. Nancy Brooks, Kieran Donaghy and Gerrit Knaap. Oxford University Press, 2011

Böhlmark, A., and A. Willén (2017) "Tipping and the Effects of Segregation" *IFAU Working Paper 2017:14*

Böhlmark, A., and M. Lindquist (2006) "Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden" *Journal of Labor Economics* 24(4): pp. 879-896

Calonico, S., M.D. Cattaneo, and R. Titiunik (2014) "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs" *Econometrica* 82(6): pp. 2295-2326

Card, D., and D. Lee (2008) "Regression discontinuity inference with specification error" *Journal of Econometrics* 142(2): pp. 655-674

Card, D., and J. Rothstein (2007) "Racial Segregation and the Black-White Test Score Gap" *Journal of Public Economics* 91(11): pp. 2158-2184

Card, D., A. Mas, and J. Rothstein (2008) "Tipping and the Dynamics of Segregation" *Quarterly Journal of Economics* 123(1): pp. 177–218

Card, D,, A. Mas, and J. Rothstein (2011) "Are Mixed neighborhoods Always Unstable? Two-Sided and One-Sided Tipping" in *Neighborhood and Life Chances: How Place Matters in Modern America*, eds. Harriet Newburger, Eugenie Birch and Susan Wachter, University of Pennsylvania Press, 2011

Charles, C. (2000) "Neighborhood Racial-Composition Preferences: Evidence from a Multiethnic Metropolis" *Social Problems* 47(3): pp. 379-407

Charles, C. (2003) "The Dynamics of Racial Residential Segregation" *Annual Review of Sociology* 29: pp. 167-207

Chetty, R., J. Friedman, N. Hilger, E. Saez, D. Schanzenbach, and D. Yagan (2011) "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star" *Quarterly Journal of Economics* 126(4): pp. 1593-1660

Chetty, R., N. Hendren, and L. Katz (2015). "The effects of exposure to better neighborhoods on children: new evidence from the moving to opportunity experiment" *Working Paper NBER 21156*

Chiswick, B. R. (1991) "Speaking, Reading, and Earnings among Low-Skilled Immigrants" *Journal of Labor Economics* 9(2): pp. 149– 170

Chiswick, B. R., and Paul W. Miller (2004) "Where Immigrants Settle in the United States" *Journal of Comparative Policy Analysis* 6(2): pp. 185-197

Chiswick, B. R., and P. W. Miller (2005) "Do Enclaves Matter in Immigrant Adjustment?" *City and Community* 4(1): pp. 5–35

Collins, W., and R. Margo (2000) "Residential segregation and socioeconomic outcomes: When did ghettos go bad?" *Economic Letters* 69(2): pp. 239-243

Cutler, D., and E. Glaeser (1997) "Are Ghettos Good Or Bad?" *Quarterly Journal of Economics* 112(3): pp. 827–872

Cutler, D., E. Glaeser, and J. Vigdor (1999). "The Rise and Decline of the American Ghetto" *Journal of Political Economy* 107 (June): pp. 455-506

Cutler, D., E. Glaeser, and J. Vigdor (2008) "When are ghettos bad? Lessons from immigrant segregation in the United States" *Journal of Urban Economics* 63(3): pp. 759–774

Damm, A.P. (2009) "Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence" *Journal of Labor Economics* 27(2): pp. 281–314

Deming, D., S. Cohodes, J. Jennings, and C. Jencks (2013) "School Accountability, Postsecondary Attainment and Earnings" *NBER Working Paper No. 19444*

DHS (2013) *Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2012*, Washington, D.C.: The Office of Immigration Statistics, U.S. Department of Homeland Security

Easterly, W. (2009) "Empirics of Strategic Interdependence: The Case of the Racial Tipping Point" *BE Journal of Macroeconomics* 9(1): Article 11

Edin, P. A., P. Fredriksson, and O. Åslund (2003) "Ethnic Enclaves and the Economic Success of Immigrants: Evidence from a Natural Experiment" *Quarterly Journal of Economics* 118(1): pp. 489–526

Farley, R. C. Steeh, M. Krysan, T. Jackson, and K. Reeves (1994) "Stereotypes and Segregation: Neighbourhoods in the Detroit Area"*American Journal of Sociology* 100(3): pp. 750–80

Fredriksson, P., B. Öckert, and H. Oosterbeek (2013) "Long-term effects of class size" *Quarterly Journal of Economics* 128(1): pp. 249–85

Glaeser, E., B. Sacerdote, and J. Scheinkman (1996) "Crime and Social Interactions" *The Quarterly Journal of Economics* 111(2): pp. 507-548

Gould, E., V. Lavy, and M. Paserman (2009) "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence" *The Economic Journal* 119: pp. 1243–1269

Gronqvist, H., S. Niknami, and P. O. Robling (2016) "Childhood exposure to segregation and long-run criminal involvement" *SOFI Working Paper No. 1/2015*

Haider, S., and G. Solon (2006) "Life-Cycle Variation in the Association between Current and Lifetime Earnings" *American Economic Review* 96(4): pp. 1308-20

Harding, D. (2003) "Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping Out and Teenage Pregnancy" *American Journal of Sociology* 109(3): pp. 676-719

Ihlanfeldt, K., and D. Sjoquist (1998) "The Spatial Mismatch Hypothesis: A Review of Recent Studies and Their Implications for Welfare Reform" *Housing Policy Debate* 9(4): pp. 849–892

Ihlanfeldt, K., and B. Scafidi (2002) "Black Self-Segregation as a Cause of Housing Segregation: Evidence from the Multi-City Study of Urban Inequality" *Journal of Urban Economics* 51(2): pp. 366-390

Jacob, B. (2004). "Public Housing, Housing Vouchers and Student Achievement: Evidence from Public Housing Demolitions in Chicago" *American Economic Review* 94(1): 233-258.
Jacob, R., and P. Zhu (2012) *A Practical Guide to Regression Discontinuity* (NY, NY)
Kahn, L. (2015). "Wage Compression and the Gender Pay Gap" *IZA World of Labor 150*
Kain, J. (1968) "Housing Segregation, Negro Employment, and Metropolitan Decentralization." *Quarterly Journal of Economics* 82: pp.175-197
Katz, L., J. Kling, and J. Liebman (2001) "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment" *Quarterly Journal of Economics,* 116(2): 607-654
Kling, J., J. Liebman, and L. Katz (2007) "Experimental Analysis of Neighborhood Effects" *Econometrica* 75(1): 83-119
Krysan, M., R. Farley, and M. Couper (2008) "In the Eye of the Beholder" *Du Bois Review* 5(1): pp. 5-26
Lazear, E. (1999) "Culture and Language" *Journal of Political Economy* 107(6): pp. 95–126
Leamer, E. (1978) *Specification Searches: Ad Hoc Inference with Non Experimental Data* (New York: John Wiley and Sons)
Lee, D. (2008) "Randomized Experiments from Non-random Selection in U.S. House Elections" *Journal of Econometrics* 142(2): pp. 675-97
Lee, D., and T. Lemieux (2010) "Regression Discontinuity Designs in Economics" *Journal of Economic Literature* 48: pp. 281-355
Le Grand, C., and R. Szulkin (2003) "Permanent Disadvantage or Gradual Integration: Explaining the Immigrant-Native Earnings Gap in Sweden" *Labour* 16(1): pp. 37-64
Lovenheim, M., and A. Willén (2016) "The Long-Run Effects of Teacher Collective Bargaining" *CESifo Working Paper Series No. 5977*
Ludwig, J., and D. Miller (2005) "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design" *NBER Working Paper 11702*
Massey, D., and N. Denton (1998). *American Apartheid: Segregation and the Making of the Underclass* (Cambridge: Harvard University press)
MPI (2010). *Largest U.S. Immigrant Groups over Time, 1960-Present* (Washington, D.C., Migration Policy Institute)
Munshi, K. (2003) "Networks in the modern economy: Mexican migrants in the U.S. labor market" *Quarterly Journal of Economics* 118: pp. 549-599
Muster, S and R. Andersson (2006). "Employment, Social Mobility and Neighborhood Effects: The Case of Sweden" *International Journal of Urban and Regional Research* 30(1): pp. 120-140
Mårdberg, B., and B. Carlstedt (1993) "Construct Validity of the Swedish Enlistment Battery" *Scandinavian Journal of Psychology* 34: pp. 353-362
OECD (2002). *Education at a Glance* (Paris: OECD)
O'Flaherty, B. (2015) *The Economics of Race in the Unites States* (Cambridge: Harvard University Press)
Oliver, E., and J. Wong (2003). "Intergroup prejudice in multiethnic settings" *American Journal of Political Science* 47(4): pp. 567-82
Ong, C. (2015). "Tipping in Dutch big city neighborhoods" *Urban Studies*: pp. 1-22
Pan, J. (2015) "Gender Segregation in Occupations: The Role of Tipping and Social Interactions" 33(2): pp. 365-408
Patacchini, E., and Y. Zenou (2012) "Ethnic Networks and Employment Outcomes" *Regional Science and Urban Economics* 42(6): pp. 938–949
Quillian, L. (2014) "Does Segregation Create Winners and Losers? Residential Segregation and Inequality in Educational Attainment" *Social Problems* 61(3): pp. 402-426

Sanbonmatsu, L., J. Kling, G. Duncan, and J. Brooks-Gunn (2007) "Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment" *Journal of Human Resources* XLI(4): pp. 649-691

Schelling, T. (1971) "Dynamics Models of Segregation" *Journal of Mathematical Sociology* 1: pp. 143-186

SOU (2011) Vård efter behov och på lika villkor – en mänsklig rättighet*, Swedish Government Official Report 2011:48*

Staiger, D. and J. Stock (1997). "Instrumental Variables Regression with Weak Instruments" *Econometrica* 65(3): pp. 557-586

Stark, O. (1991) *The Migration of Labor,* Oxford: Blackwell

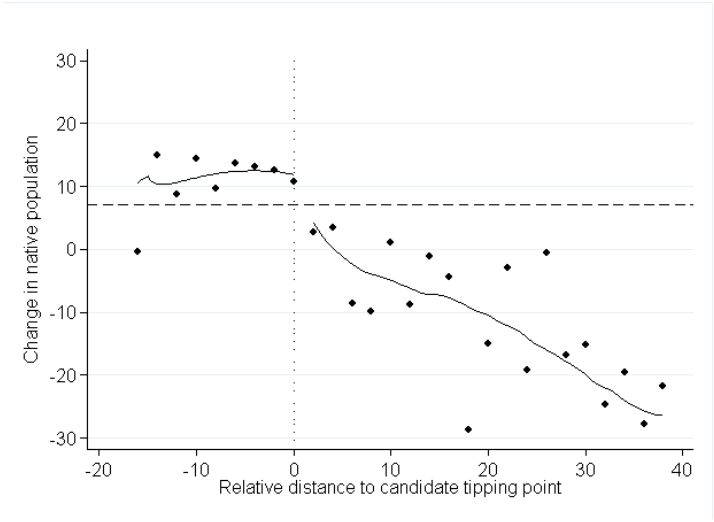
Statistics Sweden (2015) *Folkmängd efter födelseland*, accessed August 27, 2016, from: http://www.scb.se/Statistik/BE/BE0101/2015A01J/BE0101-Folkmangd-fodelseland-1900-2015.xlsx

UNHCR (2015) *Mid-Year Trends 2015* (The UN Refugee Agency), accessed June 13, 2016, from: http://www.unhcr.org/cgi-bin/texis/vtx/home/opendoc PDFVieweer.html?docid=56701b969&query=mid-2015

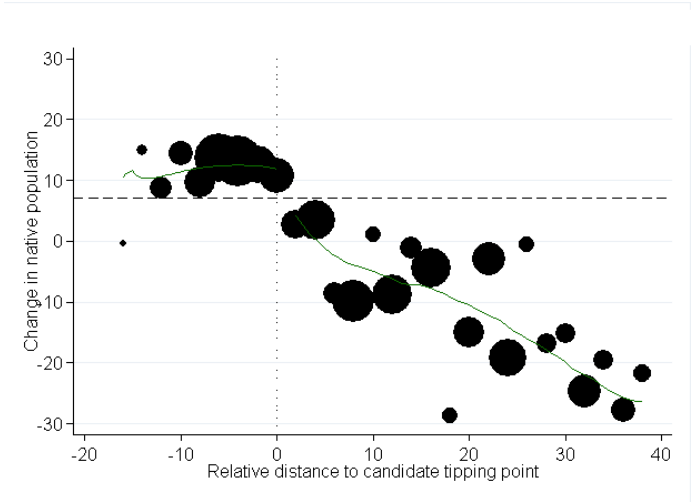
Zimmerman, K. (1996) "European Migration: Push and Pull" *International Regional Science Review* 19: pp. 95-128

Åslund O., and P. Fredriksson (2009) "Peer Effects in Welfare Dependence. Quasi-Experimental Evidence" *The Journal of Human Resources* 44(3): pp. 798–825

Åslund, Olof, P. Edin, P. Fredriksson, and H. Grönqvist (2011) "Peers, Neighborhoods, and Immigrant Student Achievement: Evidence from a Placement Policy" *American Economic Journal: Applied Economics* 3(2): pp. 67-95



(a)



(b)

Figure 2.1: Discontinuities in native population growth around candidate tipping point

Notes: Dots show mean change in neighborhood native population between 1990 and 2000 as a percentage of total neighborhood population in 1990, grouping neighborhoods into 2% bins by the deviation in immigrant share from the estimated tipping point in 1990. In Panel B, the size of each dot is weighted by the total size of the neighborhoods used to obtained that dot. The dashed horizontal lines represent the unconditional means, and the dotted vertical lines depict the estimated tipping points (normalized to zero). The solid lines are local linear regressions fit separately on either side of the tipping point weighted by the size of the neighborhoods, using an Epanechnikov kernel and a bandwidth of 4. Only the 1/3 of the sample not used for identifying the location of the tipping points is used for these visual depictions.

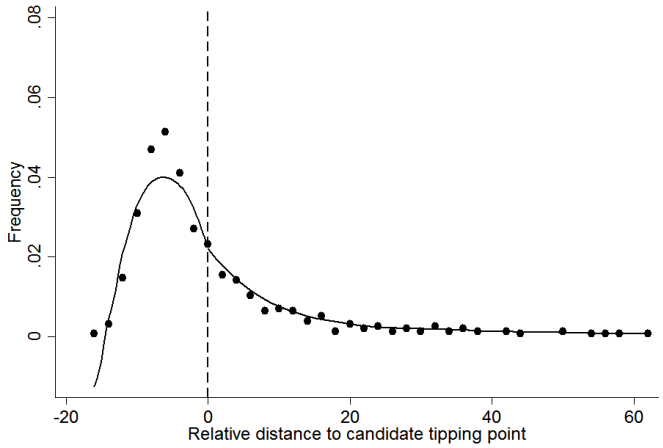
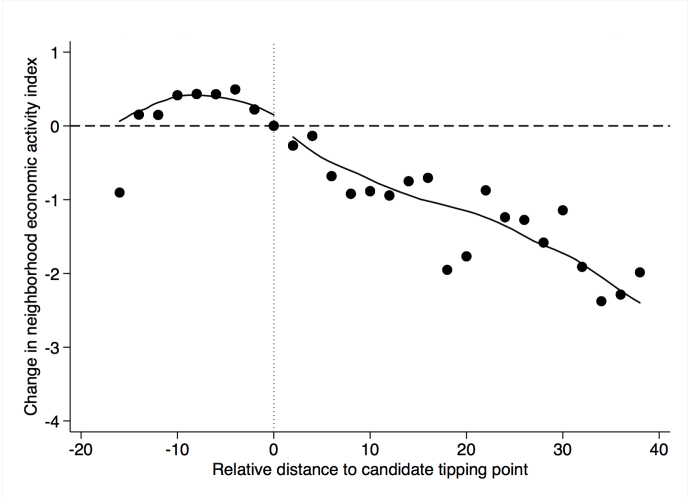
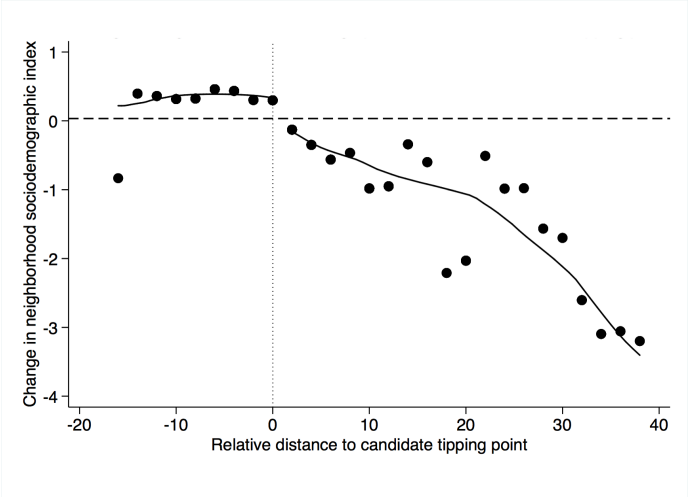


Figure 2.2: Density plot of fraction non-Western immigrants in base year
Notes: The x-axis represents the deviation in non-Western immigrant share from the estimated tipping point, grouping neighborhoods into 2% bins. The y-axis measures the frequency of observations for each of the 2% neighborhood bins. The vertical line depicts the estimated tipping point (normalized to zero). The solid line is a local linear regression fitted separately on either side of the tipping point, using an Epanechnikov kernel and a bandwidth of 4. The full sample has been used for this depiction.



(a) Economic Activity Index



(b) Sociodemographic Index

Figure 2.3: Discontinuities in neighborhood composition around candidate tipping point

Notes: Dots show mean change in neighborhood economic activity index (a) and sociodemographic index (b), grouping neighborhoods into 2% bins by the deviation in immigrant share from the estimated tipping point in 1990. The Economic Activity Index is based on three labor market variables (average employment income, average education and fraction employed) while the Sociodemographic Index is based on four sociodemographic variables (gender balance, age profile, fraction immigrants and fraction on social welfare). For each of these indices, we use unity-based normalization to bring the values of each of the individual variables into the range [0,1], take their sum, and then standardize the index to have a mean of zero and a standard deviation of one. The dashed horizontal lines represent the unconditional means, and the dotted vertical lines depict the estimated tipping points (normalized to zero). The solid lines are local linear regressions fit separately on either side of the tipping point weighted by the size of the neighborhoods, using an Epanechnikov kernel and a bandwidth of 4. Only the 1/3 of the sample not used for identifying the location of the tipping points is used for these visual depictions.

Table 2.1: Summary statistics of neighborhoods in sample

	1990	2000
Panel A: 0-5% Immigrants		
Percent Neighborhoods	2.10	1.98
Native Growth	5.53 (19.12)	2.25 (19.76)
Total Growth	7.43 (19.86)	5.69 (21.72)
Panel B: 5-20% Immigrants		
Percent Neighborhoods	73.03	65.06
Native Growth	12.54 (28.10)	7.70 (26.81)
Total Growth	16.42 (31.17)	12.64 (32.40)
Panel C: 20-40% Immigrants		
Percent Neighborhoods	19.15	21.42
Native Growth	0.27 (34.37)	-2.18 (21.99)
Total Growth	14.07 (41.73)	11.35 (25.98)
Panel D: 40-100% Immigrants		
Percent Neighborhoods	5.72	11.54
Native Growth	-12.95 (15.92)	-6.21 (15.36)
Total Growth	7.51 (19.08)	12.26 (23.59)

Notes: Authors' own calculations. The unit of observation is a neighborhood as identified by the SAMS code. Values based on unweighted means across Sweden's three meteropolitan areas.

Table 2.2: Descriptive statistics of individuals in sample

	Whole Country			Meteropolitan Areas			Analysis Sample		
	Mean	S.D.	Observations	Mean	S.D.	Observations	Mean	S.D.	Observations
Panel A: Young Cohort									
Female	0.488	0.500	1,408,511	0.486	0.500	299,094	0.480	0.500	233,091
Age	4.921	3.205	1,408,511	4.623	3.224	299,094	4.661	3.225	233,091
Mother's Education	11.530	2.142	1,126,038	11.802	2.299	287,365	11.769	2.280	225,416
Father's Education	11.623	2.333	1,103,778	12.030	2.530	278,958	11.977	2.503	218,015
Parental Income (000s SEK)	145.857	83.427	1,118,008	157.862	102.227	284,737	158.129	100.188	223,517
Panel B: Middle Cohort									
Female	0.489	0.500	1,443,459	0.486	0.500	271,548	0.486	0.500	220,671
Age	12.150	3.176	1,443,459	12.122	3.200	271,548	12.251	3.247	220,671
Mother's Education	11.447	2.242	1,074,293	11.786	2.388	255,682	11.749	2.374	210,925
Father's Education	11.573	2.450	1,039,277	12.034	2.639	242,993	11.983	2.619	199,122
Parental Income (000s SEK)	174.926	96.468	1,049,611	194.683	120.396	246,345	194.539	118.912	204,264
Panel C: Old Cohort									
Female	0.490	0.500	1,318,727	0.491	0.500	374,807	0.497	0.500	348,397
Age	37.117	3.191	1,318,727	36.983	3.205	374,807	36.989	3.203	348,397
Mother's Education	9.984	1.652	545,278	10.441	1.959	135,327	10.438	1.956	128,803
Father's Education	10.480	2.106	331,966	11.069	2.408	86,018	11.065	2.402	81,965
Parental Income (000s SEK)	109.049	93.030	297,356	131.592	109.491	76,644	131.897	109.180	73,098
Employed	0.894	0.308	1,318,727	0.873	0.333	374,807	0.889	0.314	348,397
Employment Income (000s SEK)	165.555	114.868	1,318,727	177.743	132.507	374,807	181.568	130.369	348,397
Government-funded Benefits (000s SEK)	23.608	37.976	1,318,727	24.971	39.279	374,807	24.888	38.934	348,397
Social Welfare Participation	0.069	0.254	1,318,727	0.086	0.281	374,807	0.082	0.275	348,397

Notes: The unit of observation is an individual. The first three columns display descriptive statistics of all individuals in Sweden, the second three columns display statistics of only those that resided in Stockholm, Malmo and Gothenburg in 1990, and the last three columns display descriptive statics of only those included in our analytical sample. Young Cohort refers to individuals born between 1980 and 1990, Middle Cohort refers to individuals that started school between 1980 and 1990, and Old Cohort refers to individuals born between 1948 and 1958.

Table 2.3: Regression discontinuity models for changes in population composition around the candidate tipping points

	OLS Estimation			IV Estimation		
	Native Growth	Immigrant Growth	Population Growth	Native Growth	Immigrant Growth	Population Growth
Treatment Measure	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Beyond TP	-0.091**	0.017	-0.074*	-0.142***	0.026	-0.116*
	(0.038)	(0.021)	(0.043)	(0.053)	(0.036)	(0.062)
First-stage F-statistic	-	-	-	295.45	295.45	295.45
Observations	520	520	520	520	520	520
R-Squared	0.305	0.362	0.081	0.324	0.368	0.088

Notes: The unit of observation is a neighborhood as identified by the SAMS code. The results in Columns (i) through (iii) are obtained from estimating equation (2). The results in Columns (iv) through (vi) are obtained from estimating a modified version of equation (2) in which the fraction of the decade that the neighborhood was exposed to tipping has been instrumented by whether the neighborhood was above or below the candidate tipping point in the base year. All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. The sample is the 1/3 sample not used for identifying the location of the tipping points. Demographic controls are years of schooling, income and gender, all measured in the base year. Regressions are weighted by the size of the neighborhoods. All specifications include metropolitan fixed effects. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 2.4: Testing for jumps in baseline covariates around the candidate tipping points

	Native characteristics	Immigrant characteristics
Age	0.261	0.625
	(0.322)	(0.424)
Percent Female / Immigrant Female	-0.608	0.083
	(0.853)	(0.167)
Years of Schooling	-0.194	0.024
	(0.200)	(0.216)
Social Welfare	-0.004	0.041**
	(0.008)	(0.018)
Social Insurance Benefits	-0.886	-2.868
	(1.041)	(2.032)
Income	-1.954	-3.302
	(4.271)	(5.117)
Employment	0.002	-0.017
	(0.006)	(0.019)
Compulsory School Drop-Out	0.013	-0.007
	(0.010)	(0.016)
High School Drop-Out	-0.024	0.007
	(0.022)	(0.022)
University Enrollment	-0.017	-0.005
	(0.015)	(0.011)
Number of Children	0.007	0.057
	(0.062)	(0.050)

Notes: Each row is a seperate estimation of equation (2) stratified by nativity status. All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. Sample is the 1/3 sample not used for identifying the location of the tipping points. Regressions are weighted by the size of the neighborhoods. All models include metropolitan fixed effects. Dependent variables are measured in the base year. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 2.5: The reduced form effect of neighborhood composition on educational attaiment

	Compulsory School					High School		Higher Education	
	GPA	Math	English	Swedish	Science	GPA	Academic Track	University Enrollment	Years of Schooling
Panel A: Immigrants									
I. Young Cohort									
Beyond TP	1.316	0.555	1.378	1.798	-0.019	0.453	-0.007	0.007	0.031
	(1.604)	(1.386)	(1.683)	(1.862)	(0.031)	(1.502)	(0.031)	(0.027)	(0.095)
Percentage Change	2.789	1.229	2.731	3.880	-8.260	0.975	-1.048	2.593	0.241
Observations	9,472	9,286	9,248	7,871	4,381	4,380	4,381	9,675	9,675
II. Middle Cohort									
Beyond TP	-0.204	-1.976*	-3.246*	-0.401	0.003	-0.271	-0.041	0.002	0.140
	(1.219)	(1.161)	(1.779)	(2.042)	(0.020)	(1.883)	(0.029)	(0.032)	(0.151)
Percentage Change	-0.429	-4.362	-6.360	-0.874	-1.251	-0.594	-6.557	0.870	1.115
Observations	9,191	8,906	8,415	8,961	6,477	3,257	6,477	9,441	9,441
Panel B: Natives									
I. Young Cohort									
Beyond TP	-2.241**	-1.395	-0.967	-1.799**	-0.008	-1.377**	-0.030	-0.013	-0.074
	(1.027)	(0.856)	(0.841)	(0.877)	(0.016)	(0.585)	(0.019)	(0.011)	(0.049)
Percentage Change	-4.231	-2.687	-1.811	-3.415	-3.545	-2.623	-4.918	-4.815	-0.592
Observations	63,907	62,725	62,704	62,516	32,266	32,245	32,266	63,676	63,676
II. Middle Cohort									
Beyond TP	-1.176*	-0.202	-2.237***	-1.861**	0.002	-0.118	-0.014	0.004	-0.026
	(0.681)	(0.984)	(0.820)	(0.805)	(0.011)	(0.770)	(0.016)	(0.012)	(0.052)
Percentage Change	-2.286	-0.406	-4.231	-3.618	0.952	-0.231	-2.692	-1.667	-0.202
Observations	60,145	59,298	59,193	59,274	48,399	23,605	48,399	59,490	59,490

Notes: The unit of observation is an individual that resided in one of the 520 neighborhoods not used for identifying the location of the tipping points. The results are obtained from estimating equation (3). All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. Demographic controls are gender, mother's education, father's education, parental income and binaries for whether this information was not available for the individual, all measured in the base year. For the Old Cohort, parental education and income have been replaced with own education and income. All models include birth year and metropolitan area fixed effects. Immigrants refer to individuals born in, or that have at least one parent born in, a non-Western country. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 2.6: The reduced form effect of neighborhood composition on labor market outcomes

	Intensive Margin			Extensive Margin		
	Self-Employment Income	Employment Income	Gov.-Funded Benefits	Self-Employment Income	Employment Income	Gov.-Funded Benefits
Panel A: Immigrants						
I. Young Cohort						
Beyond TP	0.098	0.285	0.250	0.008	0.016	0.025
	(0.268)	(0.103)	(0.232)	(0.010)	(0.019)	(0.023)
Percentage Change	0.948	2.452	2.518	25.807	2.025	16.667
Observations	391	8,043	1,723	10,266	10,266	10,266
II. Middle Cohort						
Beyond TP	0.131	0.120	0.214	0.014	0.006	0.019
	(0.164)	(0.247)	(0.246)	(0.014)	(0.020)	(0.023)
Percentage Change	1.230	0.986	2.202	16.981	0.779	5.429
Observations	626	7,955	3,867	10,307	10,307	10,307
III. Old Cohort						
Beyond TP	-0.024	0.384	-0.102	0.004	0.032	-0.008
	(0.052)	(0.281)	(0.068)	(0.005)	(0.022)	(0.007)
Percentage Change	-0.215	3.112	-1.050	2.857	1.786	-7.080
Observations	1,429	13,767	2,544	23,253	23,253	23,253
Panel B: Natives						
I. Young Cohort						
Beyond TP	0.003	-0.029	-0.001	0.001	-0.002	-0.000
	(0.025)	(0.101)	(0.132)	(0.003)	(0.008)	(0.013)
Percentage Change	0.030	-0.246	0.010	3.334	-0.235	-0.000
Observations	2,084	57,066	11,517	65,982	65,982	65,982
II. Middle Cohort						
Beyond TP	0.016	0.051	0.009	0.002	0.005	0.002
	(0.042)	(0.091)	(0.098)	(0.004)	(0.007)	(0.010)
Percentage Change	0.153	0.412	0.092	3.965	-0.555	0.435
Observations	3,446	54,544	28,439	62,476	62,476	62,476
III. Old Cohort						
Beyond TP	0.027	-0.030	-0.034	0.002	-0.002	-0.003
	(0.070)	(0.130)	(0.040)	(0.007)	(0.027)	(0.004)
Percentage Change	0.249	0.240	-0.353	4.002	0.247	-2.498
Observations	6,532	74,488	10.746	93.953	93.953	93.953

Notes: The unit of observation is an individual that resided in one of the 520 neighborhoods not used for identifying the location of the tipping points. The results are obtained from estimating equation (3). All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. Demographic controls are gender, mother's education, father's education, parental income and binaries for whether this information was not available for the individual, all measured in the base year. For the Old Cohort, parental education and income have been replaced with own education and income. All models include birth year and metropolitan area fixed effects. All dependent variables are measured in 2011. Immigrants refer to individuals born in, or that have at least one parent born in, a non-Western country. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 2.7: The reduced form effect of neighborhood composition on stayers and leavers

	9th grade GPA	9th grade Swedish	9th grade math	9th grade English	High school Academic track	High school science program	High school GPA
Panel A: Native stayers							
Young Cohort	-1.018	-1.356	-0.532	-1.171	-0.010	-0.004	-0.876
	(1.457)	(1.242)	(1.180)	(1.018)	(0.023)	(0.020)	(0.767)
Middle Cohort	-0.830	-1.324	0.041	-2.650**	0.019	-0.001	0.292
	(1.207)	(1.205)	(1.436)	(1.155)	(0.025)	(0.019)	(0.975)
Panel B: Native leavers							
Young Cohort	-3.109***	-2.128**	-1.195	-1.702*	-0.046**	-0.013	-1.835***
	(1.126)	(0.913)	(1.043)	(0.974)	(0.021)	(0.023)	(0.659)
Middle Cohort	-1.405*	-1.934**	0.036	-2.227**	-0.023	0.005	-0.060
	(0.788)	(0.913)	(1.207)	(0.843)	(0.017)	(0.015)	(0.890)
Panel C: Immigrant stayers							
Young Cohort	2.642	2.288	1.196	2.408	0.044	-0.019	-0.974
	(2.246)	(2.306)	(2.036)	(2.031)	(0.037)	(0.032)	(1.963)
Middle Cohort	0.453	-0.127	-1.987	-1.017	0.010	-0.032	-3.434
	(1.946)	(2.407)	(1.578)	(2.107)	(0.035)	(0.026)	(2.227)
Panel D: Immigrant leavers							
Young Cohort	1.210	1.674	0.834	1.079	-0.017	-0.000	2.077
	(1.535)	(1.776)	(1.395)	(1.978)	(0.036)	(0.042)	(1.794)
Middle Cohort	-0.455	-0.586	-2.036	-4.396**	-0.063**	0.025	2.640
	(1.401)	(2.087)	(1.584)	(1.965)	(0.028)	(0.021)	(1.960)

Notes: The unit of observation is an individual that resided in one of the 520 neighborhoods not used for identifying the location of the tipping points. The results are obtained from estimating equation (3). All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. Demographic controls are gender, mother's education, father's education, parental income and binaries for whether this information was not available for the individual, all measured in the base year. For the Old Cohort, parental education and income have been replaced with own education and income. All models include birth year and metropolitan area fixed effects. Natives refer to individuals not born in, and that do not have a parent born in, a non-Western country. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

CHAPTER 3

The Long-run Effects of Teacher Collective Bargaining[a]

Abstract

Teacher collective bargaining is a highly debated feature of the education system in the US. This paper presents the first analysis of the effect of teacher collective bargaining laws on long-run labor market and educational attainment outcomes, exploiting the timing of passage of duty-to-bargain laws across cohorts within states and across states over time. Using American Community Survey data linked to each respondent's state of birth, we examine labor market outcomes and educational attainment for 35-49 year olds, separately by gender. We find robust evidence that exposure to teacher collective bargaining laws worsens the future labor market outcomes of men: in the first 10 years after passage of a duty-to-bargain law, male earnings decline by $1,974 (or 3.64%) per year and hours worked decrease by 0.43 hours per week. The earnings estimates for men indicate that teacher collective bargaining reduces earnings by $198.1 billion in the US annually. We also find evidence of lower male employment rates, which is driven by lower labor force participation. Exposure to collective bargaining laws leads to reductions in the skill levels of the occupations into which male workers sort as well. Effects are largest among black and Hispanic men. Estimates among women are often confounded by secular trend variation, though we do find suggestive evidence of negative impacts among non-white women. Using data from the 1979 National Longitudinal Survey of Youth, we demonstrate that collective bargaining laws lead to reductions in measured non-cognitive skills among young men.

[a] We are grateful to David Autor, Dan Black, Maria Fitzpatrick, Richard Freeman, Steve Rivkin, Tim Sass, Mark Steinmeyer and seminar participants at the 2015 Association for Education Finance and Policy annual meeting, the CESifo Economics of Education Conference, Southern Methodist University, the Federal Reserve Bank of Cleveland, and the University of Mississippi for helpful comments.

1. Introduction

Teacher collective bargaining is a prevalent and contentious feature of the US education system. Over 60% of teachers in the United States currently are covered by a collectively-bargained contract (Frandsen 2016), and recently many states have weakened the ability of teachers' unions to negotiate contracts. For example, in 2011 Wisconsin, Indiana, Idaho and Tennessee passed legislation that greatly reduced the scope of teacher bargaining with school districts, and in 2012 Michigan passed a public employee right-to-work law that sought to limit teacher union negotiating power. In 2014, the ruling in *Vergara v. California* argued that the tenure and teacher retention policies that are a main focus of collective bargaining violated the constitutionally-guaranteed right to an adequate education for each child in California.[2] These court rulings and legislative actions have reignited a debate over the proper role of teacher collective bargaining in the US education system. One of the core factors on which this debate rests is how teacher collective bargaining impacts student outcomes. Despite the large amount of policy attention directed toward the role of teachers' unions in education, there is a lack of empirical research that credibly and comprehensively addresses this question.

A central hurdle facing the prior teachers' union literature is the lack of student outcome data linked to exogenous variation in teacher collective bargaining. Much of the cross-sectional variation in teacher bargaining is driven by state public sector union laws that determine the obligations of school districts to negotiate with teachers. These laws were passed in the 1960s-1980s, when only sparse data were available on student outcomes that could be matched to one's school district. The small set of studies that have examined the relationship between teacher collective bargaining and student outcomes from this time period have used high school dropout rates (Hoxby 1996; Lovenheim 2009) or state-level SAT scores (Kleiner and Petree 1988). These analyses reach different conclusions, and their focus on a

[2] This ruling was reversed in 2016 by the California Court of Appeals, and the reversal was subsequently upheld by the California Supreme Court.

limited set of performance measures does not yield a complete picture of the effects of teacher collective bargaining on students. More recent studies have better student achievement data but lack exogenous variation in teacher collective bargaining (e.g., Lott and Kenny 2013; Strunk 2011; Moe 2009).

In this paper, we present the first evidence in the literature on how teacher collective bargaining laws affect long-run outcomes of students. We focus on duty-to-bargain (DTB) laws, which require districts to negotiate with teachers' unions in good faith. Prior work has shown extensive evidence that these laws increase union membership and the probability that a district elects a union to bargain collectively (Frandsen 2016; Lovenheim 2009; Hoxby 1996; Saltzman 1985). We use the timing of the passage of DTB laws, which occurred between 1960 and 1987 (see Figure 1), linked with educational and labor market outcomes among 35-49 year olds in the 2005-2012 American Community Survey (ACS), to provide new evidence on how teacher collective bargaining impacts a broad array of long-run outcomes.

Critical to our identification strategy is the ability to link ACS respondents to their state of birth, which allows us to account for any endogenous migration of families across states with different collective bargaining laws. We employ cross-cohort difference-in-difference event study models that examine how outcomes changed among students who were differentially exposed to duty-to-bargain laws that had been in place for different lengths of time based on what state and in what year they were born. The sources of variation we exploit come from within-state changes in outcomes across birth cohorts as a function of time since passage of a DTB law and cross-state differences in the timing of when (or whether) these laws were passed.

Identification is complicated by the potential for other policies, secular trends, and unobserved shocks to affect the outcomes of interest. We show extensive evidence that our estimates are not being driven by such factors. First, our models include controls for other

important policies during this period to which students may have been exposed. Second, we explicitly test for the existence of pre-treatment trends in outcomes across cohorts. Third, the results are robust to directly controlling for pre-treatment trends. Fourth, our results are not being driven by the general union environment in the state, are not influenced by the urbanicity of the population, are not correlated with the prevalence of social unrest in the state when our sample was of school age, are not influenced by the political environment in the state, and are robust to accounting for region-specific cohort shocks. Fifth, we perform permutation tests in which we randomly assign the year of duty-to-bargain law passage across states. Finally, our estimates are not biased by cross-state mobility of those with school-age children. Taken together, these results provide extensive evidence that supports the causal interpretation of our estimates.

Women's educational and labor market outcomes were subject to strong secular changes among the cohorts we examine (Goldin, Katz and Kuziemko 2006; Blau and Kahn 2013; Bick and Bruggeman 2014), and we thus analyze outcomes separately by gender.[3] Among men, our estimates point to negative effects of exposure to teacher collective bargaining laws on the long-run labor market outcomes of students. These results are consistent with the "rent-seeking" hypothesis of teacher unionization (Hoxby 1996).[4] We present estimates of DTB law exposure at 5, 10, and 15 years post law passage, and we focus on the 10-year estimates because they show the effects among the first cohort that spent nearly the entirety of its schooling years in a collective bargaining environment.

At 10 years of DTB exposure, male annual earnings decline by $1,973.48 (or 3.64%) and weekly hours worked are reduced by 0.43 (or 1.09%). These individuals are also 0.9 of a percentage point less likely to be employed, are 0.8 of a percentage point less likely to be in

[3] These secular trends reflect reduced gender-based discrimination, rising expectations of future labor market participation among women, increased female collegiate attendance, and expanded female labor market opportunities.
[4] The rent-seeking hypothesis of teachers' unions states that unions lead to a re-allocation of resources towards teachers while also making educational resources less productive. See Section 2 for a more in-depth discussion.

the labor force, and sort into lower-skilled occupations. However, collective bargaining laws have only a modest effect on educational attainment. Our estimates therefore suggest that the effect of teacher collective bargaining on labor market outcomes is driven by declines in human capital accumulation that do not show up in years of education. This finding motivates our analysis using the 1979 National Longitudinal Survey of Youth (NLSY) that shows declines in non-cognitive skills due to collective bargaining exposure.

We further demonstrate that the negative effects of duty-to-bargain laws are particularly pronounced among black and Hispanic males: annual earnings decline by $3,019 (8.77%), hours worked per week decline by 0.58 (1.76%), the likelihood of being employed is 0.9 of a percentage point lower, and years of schooling and occupational skill are significantly lower at 10 years of exposure. Collective bargaining laws also lead to worse labor market outcomes among white and Asian men, but the effects are more modest in magnitude.

We do not find consistent effects of collective bargaining law exposure on female labor market and education outcomes. Most of the point estimates are negative, but they are much smaller than those for men. Further, they show clear evidence of differential pre-treatment trends. Thus, our empirical approach does not appear valid for women; we cannot draw strong conclusions about how duty-to-bargain laws affect long-run female outcomes with our approach. Importantly, there is no evidence that the secular trends for women produce similar trends among men that would threaten our identification strategy. We do find more evidence of negative effects among black and Hispanic women, which together with the male estimates suggests DTB laws predominantly affect long-run outcomes among minorities.

A drawback of our setting, common to most studies on long-run program effects, is that we cannot examine the mechanisms through which our results operate. However, when examining education policies, we ultimately care about how they impact school quality and the long-run outcomes of students, which we speak to directly in this paper. That the data do

not exist to examine all mechanisms at work in determining long-run effects further augments the importance of directly estimating impacts on these long-run outcomes themselves.

Though we are unable to comprehensively examine the mechanisms, we show that DTB laws are associated with higher expenditures on teachers and administrators but do not alter total expenditures or teacher-student ratios. Prior research also has found evidence that DTB laws reduce hours worked among teachers (Frandsen 2016) and that reduced bargaining power leads to lower fringe benefits among teachers (Litten 2017). Given the impossibility of exploring a comprehensive set of mechanisms, the long-run estimates we produce represent new evidence on the impact of DTB laws that are very important because of the prevalence of these laws, the contentiousness surrounding them, the recent rise in policies aimed at curbing teacher collective bargaining rights, and the paucity of evidence on how they affect students.

We also provide supporting estimates showing how DTB laws impact non-cognitive skills among high school students using the NLSY79. These outcomes indicate whether the long-run effects we identify are reflective of changes in human capital. Consistent with the labor market effects, we find that duty-to-bargain law exposure reduces these outcomes among men. The impacts on non-cognitive measures helps reconcile the fact that we do not see a strong educational attainment effect despite large reductions in labor market outcomes, since non-cognitive skills are likely to affect labor market outcomes more than education outcomes (Heckman, Stixrud and Urzua 2006; Heckman and Kautz 2012; Heckman, Pinto and Savelyev 2013). These estimates support our long-run findings and indicate that teacher collective bargaining laws reduce the quality of education students receive.

Taken together, our results suggest that public sector collective bargaining laws for teachers have a negative effect on male long-run labor market outcomes: decreasing male earnings in the 33 duty-to-bargain law states by 3.64% amounts to $198.1 billion of lost earnings per year. We underscore that these estimates are from a time period in which the

education system was different along many dimensions from today, so caution should be exercised in extrapolating the results to the current education system.

2. Teacher Collective Bargaining in the US

2.1. Duty-to-Bargain Laws

Prior to 1960, teachers unions in the US were predominantly professional organizations that had little role in the negotiation of contracts between teachers and school districts. Collective bargaining occurred in only a handful of large, urban school districts (e.g., New York and Detroit). Beginning with Wisconsin in 1960, states began passing public sector "duty-to-bargain" (DTB) laws, which mandated that districts have to negotiate in good faith with a union that has been elected for the purposes of collective bargaining. These laws gave considerable power to teachers' unions in the collective bargaining process. As a result, duty-to-bargain laws led to a sharp rise in teacher unionization and in the prevalence of collectively-bargained contracts (Lovenheim 2009; Saltzman 1985). In states that pass a DTB law, the vast majority of school districts elect a union for the purpose of collective bargaining, and these unions achieve contracts at very high rates (Lovenheim 2009). Thus, passage of a DTB law leads to a high fraction of teachers being covered by a collectively-bargaining contract over a short period of time.

Between 1960 and 1987, 33 states passed DTB laws, as shown in Figure 1. Most of these laws were implemented between the late-60s and late-70s. Table 1 shows the year of passage for each state as well as the set of states without such a law.[5] Of the 17 non-DTB states, 10 allow teachers to collectively bargain if both sides agree to do so. Four states (Alabama, Georgia, North Carolina, and Virginia) have no state law governing teacher collective bargaining, while three states (Mississippi, Missouri and Wyoming) outlaw

[5] Note that Washington, DC is excluded both from Table 1 and from our analysis.

collective bargaining. The states that have more restrictive laws tend to be located in the South and the West, which highlights the fact that these laws are not randomly assigned.

The focus of this paper is on how the passage of public-sector DTB laws affects the long-run outcomes of students who attended elementary or secondary schools in those states. We examine duty-to-bargain laws because these laws led to larger increases in unionization and collective bargaining rates than did the other forms of union laws (Frandsen 2016): non-duty-to-bargain union laws do not explicitly require districts to recognize unions and bargain in good faith, thus allowing them to simply refuse to engage in collective bargaining.[6]

2.2. Theoretical Predictions

One of the main ways in which DTB laws affect students is by increasing the rate and substance of bargaining between teachers and school districts. Changes in collective bargaining, in turn, can impact students through three main channels: by altering the inputs to education production, by affecting teacher effort (and thus effectiveness), and by changing the composition of teachers. The third mechanism in particular implies that the long-run effects may be larger than the short-run effects, as it takes time to alter teacher composition.

Models of public sector union behavior provide ambiguous predictions about how teacher collective bargaining should affect students. The "rent-seeking" model argues that by distorting the allocation of resources towards teachers, student outcomes may decline. The key predictions of this model are that teacher collective bargaining should lead to increases in resources going to teachers and to reductions in the returns to those resources: the resource changes induced by teachers unions reduce the efficiency of educational inputs, which negatively impacts students. By protecting teachers from being fired, unions also can reduce teacher effort and lower the quality of the teacher workforce, which will lead to worse student

[6] Our results are similar (though somewhat attenuated) when we use a more expansive definition of collective bargaining laws that includes the 10 states that allow but do not require districts to negotiate with teachers unions.

outcomes. Under the rent-seeking model, the decline effectiveness of teacher-related resources can produce worse student outcomes.[7]

In contrast to the rent-seeking model, there are several arguments suggesting that teachers unions can improve educational outcomes. Empowering teachers may result in higher achievement from a more efficient resource allocation, since educational administrators do not have full knowledge of the education production function. There also could be a "union voice" effect, whereby giving teachers a voice with which to influence their working environment makes them more productive (Freeman 1980; Gunderson 2005). A more favorable working environment could further induce more-productive workers to enter teaching.

All models of union behavior predict that teachers unions will alter district resource allocations; just examining how unions affect education inputs such as teacher pay, employment and per-student spending will not allow one to distinguish between them.[8] Where the models differ is in their predictions of the direction of any effects on student outcomes. The theoretical ambiguities highlighted above underscore the importance of conducting an empirical investigation on how teacher collective bargaining affects student outcomes.

Duty-to-bargain laws also can affect outcomes through mechanisms other than unionization, per se. Teachers unions engage in statewide advocacy that can influence non-unionized districts, and there can be union threat effects (Farber 2003) that make non-unionized districts behave like unionized ones to stave off a union vote.

2.3. Prior Research on Teacher Unionization and Collective Bargaining

The majority of research on teachers unions focuses on resource allocation effects. Collective bargaining can influence several dimensions of school resource allocation decisions: teachers

[7] The rent-seeking model does not guarantee that unionization will lead to lower student achievement. The reason is that unionization could increase total resources while also making those resources less effective. The net effect on student outcomes thus is ambiguous.

[8] It also is impossible to observe all educational inputs in most datasets. Thus, only examining the effect of unions on measured resources provides a somewhat limited description of their effect on schools and students.

typically negotiate over wage schedules, hiring and firing policies, health care and retirement benefits, work rules detailing the hours they are required to be at work and to teach, class assignments, class sizes and non-teaching duties (West 2015; Moe 2009; Strunk 2009). Research examining the effect of teacher collective bargaining on district resources has found mixed results, although data constrains have only allowed an examination of a small subset of education inputs. Studies that have exploited the rollout of DTB laws have either found positive effects on teacher salaries and per-student expenditures (Hoxby 1996) or no effects (Lovenheim 2009; Frandsen 2016).[9] Recent evidence exploiting the substantial restrictions on collective bargaining rights in Wisconsin in 2011 finds increases in teacher wage dispersion and exit (Biasi 2017; Roth 2017) as well as modest effects on average wages but a sizable impact on non-wage compensation (Litten 2017). Results from the 2011 ban on teacher collective bargaining in Tennessee indicates a reduction in teacher compensation in the form of wages and health care and shrinkage in the size of the teacher workforce (Quinby 2017).

Of first-order importance in the policy debate over the role of teachers unions in education is how collective bargaining affects student outcomes. The effects on resource allocation discussed above yield ambiguous predictions for effects on students. There currently is only a small literature on the effect of teachers' unions on academic achievement. None of these studies estimate the effect of collective bargaining on long-run labor market and educational attainment outcomes, which may differ from any short-run impacts (Ludwig and Miller 2007; Chetty et al. 2011; Deming et al. 2013; Cohodes et al. 2016). One central reason for this lack of existing work is data constraints: the teacher unionization movement took hold before consistent measures of student outcomes were collected. Thus, researchers are forced either to use a small set of outcomes from older data during the period of DTB law passage or

[9] An earlier body of work finds mixed evidence on how unions affect teacher pay. Balfour (1974), Zuelke and Frohreich (1977), and Kleiner and Petree (1988) find no effect. Eberts and Stone (1986), Moore and Raisian (1987) and Baugh and Stone (1982) find evidence of a union wage premium ranging from 3%-12%. These studies typically lack plausibly-exogenous variation in union status. See Cowen and Strunk (2015) for a review of this literature.

to use data from more recent time periods that lack exogenous variation in collective bargaining across schools.

Hoxby (1996) and Lovenheim (2009) both use the passage of duty-to-bargain laws to estimate how teacher collective bargaining affects contemporaneous high school dropout rates. Hoxby finds that collective bargaining laws lead to an increase in high school dropout rates, which is consistent with the rent-seeking model of union behavior.[10] Using an alternative unionization measure and a smaller set of states, Lovenheim (2009) finds no such effect.[11]

Much of the literature that uses more recent data to examine how unions and collective bargaining affect test scores focuses on measures of contract restrictiveness or union power. Lott and Kenny (2013) show that states with higher union dues and union expenditures have lower 4th grade proficiency rates. Strunk (2011) shows that contract restrictiveness is negatively correlated with test score levels but not with test score growth. The cross-sectional nature of these comparisons make it unlikely that they isolate the causal effect of union strength on student outcomes, as districts with strong unions tend to be in more urban, lower-income areas. Moe (2009) examines how changes over time in union contract restrictiveness within school districts in California relate to changes in student test scores. While he finds that districts with contracts that become more restrictive experience declines in test score growth, it is unlikely that within-district variation in restrictiveness over time is exogenous.[12]

Our contribution to this literature is to estimate how teacher collective bargaining affects long-run educational and labor market outcomes using an identification strategy that incorporates exogenous variation in the prevalence of collective bargaining in the state. By linking adults in different birth cohorts to their birth state, we exploit timing differences in the

[10] In contrast, Eberts and Stone (1986, 1987) find that teachers' unions increase school productivity. However, they lack exogenous variation in union status across schools, which complicates the interpretation of their results.
[11] Some prior work examines the link between teachers' unions and student outcomes using student test score data, but it typically lacks exogenous variation in union status (e.g., Kleiner and Petree 1988; Eberts and Stone 1987).
[12] Evidence from how Wisconsin's collective bargaining changes (Act 10) affected student outcomes are mixed. Biasi (2017) and Roth (2017) find increases in student test scores, while Baron (2017) finds large declines.

passage of DTB laws combined with variation in whether states ever pass such a law to overcome the identification problems and data limitations faced by prior research. Our results therefore provide the first comprehensive analysis of the causal effect of teacher collective bargaining on student outcomes, which is of first-order importance given the prevalence of teachers unions and the ongoing policy debate about their proper role in education.

3. Data

The collective bargaining data we use come from the NBER collective bargaining law dataset (Valletta and Freeman 1988).[13] These data contain, for each state and year since 1955, collective bargaining laws for each type of public sector worker. We use the laws for teachers to create an indicator variable for whether a DTB law was in place in each state and year.

We combine the collective bargaining information with 2005-2012 American Community Survey (ACS) data on individuals aged 35-49. Individuals within this age span typically have completed their education and are on a flat part of their lifetime earnings profile (Haider and Solon 2006). We observe individuals of each age in each of the eight survey years, leading to a balanced panel of age observations in our data. We construct birth cohorts by subtracting age from calendar year, and we assume each respondent begins school at the age in which his assigned birth cohort turns 6.[14] The birth cohorts range from 1956 to 1977 and correspond to students who would have been in school from 1962 (when the 1956 birth cohort was 6) to 1995 (when the 1977 birth cohort was 18). These schooling years correspond with the large rise in duty-to-bargain laws across states in the US shown in Figure 1.

A main advantages of using the ACS is the ability to link adults to their state of birth,

[13] These data are available at http://www.nber.org/publaw/.

[14] These assumptions lead to some measurement error in treatment assignment because the ACS is conducted each month and states have different school-age cutoff dates. Using the school-age cutoff dates that prevailed in 1988 (Bedard and Dhuey 2012) and assuming that ACS survey month and birth month are evenly distributed over the year, we calculate about 27% of the sample will enroll in school the year prior to their assigned birth cohort. This is likely to bias our estimates towards zero by generating changes in outcomes in the cohort just prior to DTB passage.

because collective bargaining laws might cause families to migrate across states. These laws also may cause post-schooling migration patterns to differ, as obtaining more or less skill when young could affect one's access to a more national labor market. Using each respondent's state of birth eliminates any problems associated with endogenous mobility. Of course, families can move across states such that one's state of birth differs from the state in which he or she attended school. In Section 5.5, we show that any bias resulting from such mobility is small. We also do not find evidence that parents are endogenously moving in response to DTB laws prior to a child's birth using changes in the observed composition of those born in a given state and cohort.

Because one's state of birth and birth cohort determine one's exposure to a duty-to-bargain law, we collapse the data to the state-of-birth, year-of-birth, calendar year level. Aggregation to this level is sensible because the effect of duty-to-bargain laws on student outcomes is not necessarily limited to unionized districts: these laws can impact all districts in a state through spillover and union threat effects (Farber 2003). The spillover effects come in part from union political activities that can impact educational resources and policies in all schools in the state. Additionally, union threat effects can cause non-unionized districts to begin behaving like unionized ones in order to stave off a unionization vote.

The ACS contains detailed information on educational attainment and labor market outcomes. Descriptive statistics of the variables we use are shown in Appendix Table C-1.[15] For educational attainment, we construct a *years of education* variable. In the 2008-2012 ACS, years of completed schooling are reported directly. In the 2005-2007 ACS waves, we use completed schooling levels to construct this variable.[16] We also use the ACS measures of

[15] Descriptive statistics by gender and race/ethnicity are shown in Appendix Table C-2.
[16] We code educational attainment as follows: 0 for no school completion, 4 for fourth grade completion, 6 for 5th or 6th grade completion, 8 for 7th or 8th grade completion, 9-11 for 9th through 11th grade completion, 12 for 12th grade completion and less than 1 year of college, 13 for one or more years of college with no degree, 14 for an AA degree, 16 for a BA degree, 18 for a master's or professional school degree, and 21 for a doctoral degree.

whether an individual is currently employed, unemployed or not in the labor force, as well as labor income in the previous year and hours worked per week. Labor income is the sum of wage, salary, and self-employed income over the past 12 months. Both income and hours worked are set to zero for those who do not report any income or working activity.

Finally, we construct a measure of occupational skill. Using the 2005-2012 ACS, we calculate the proportion of workers in each 4-digit occupation code that has more than a high school degree (i.e., at least some collegiate attainment). This allows us to rank occupations by the skill level of those who engage in the occupation in order to examine whether exposure to teacher collective bargaining leads workers to sort into lower- or higher-skilled occupations.

4. Empirical Methodology

We exploit within-state, cross-cohort differences in exposure to DTB laws over time driven by variation in the year of law passage combined with cross-state variation in the timing of when or whether states passed these laws in difference-in-difference framework. The effect of collective bargaining laws on student achievement is likely to vary across cohorts for two reasons. The first is that some cohorts are only exposed for part of their schooling years, which can generate time-varying treatment effects based on the length of exposure to collective bargaining laws across cohorts. The second factor that influences the time pattern of treatment effects is that the laws themselves may have time-varying impacts on resource allocation (see Lovenheim (2009) and Appendix Table C-9), the composition of teachers, and teacher effort from unions becoming more powerful or effective over time. There also can be immediate impacts of DTB law passage on student outcomes. Thus, our main empirical approach is to estimate event study models separately for men and women that allow us to non-parametrically identify time-varying treatment effects:

$$Y_{sct} = \beta_0 + \pi_{-11} I(C - t_0 + 18 \leq -11)_{sc} + \sum_{\tau=-10}^{20} \pi_\tau I(C - t_0 + 18 = \tau)_{sc} +$$
$$\pi_{21} I(C - t_0 + 18 \geq 21)_{sc} + \gamma X_{sct} + \delta_{ct} + \theta_s + \phi_t + \varepsilon_{sct}, \quad (1)$$

where Y_{sct} is one of the educational or labor market outcomes listed above for those born in state *s* in birth cohort *c* and in ACS calendar year *t*. Regressions are weighted by the number of observations that underlie each birth year-birth state-calendar year-gender cell, and all standard errors are clustered at the birth state level.

The variable $(C - t_0 + 18)$ is equal to the number of years of exposure a cohort has had to a DTB law, with *C* being the birth year and t_0 being the year of passage of the duty-to bargain law. For example, a cohort that is 19 when a duty-to-bargain law is passed will have an exposure time of -1, while a cohort that is 10 when it passes will have an exposure time of 8. This variable takes on a value of zero in states that have never had a duty-to-bargain law.[17] Hence, $I(C - t_0 + 18 = \tau)$ are indicator variables equal to 1 for each relative year to passage of a duty-to-bargain law between -10 and 20. We also include an indicator for whether time relative to a DTB law is less than or equal to -11 and for whether it is greater than or equal to 21.[18] The π_τ coefficients non-parametrically trace out pre-treatment relative trends (for π_{-11} to π_{-1}) as well as time-varying treatment effects (π_0 to π_{21}). In practice, we omit $I(C - t_0 + 18 = -1)$ such that all π estimates are relative to the year prior to DTB passage.

Equation (1) also includes birth cohort-by-calendar year (δ_{ct}), birth state (θ_s), and calendar year (ϕ_t) fixed effects. The birth cohort-by-year fixed effects are identical to age fixed effects, because birth cohort and calendar year perfectly define age. The cohort-year fixed effects control for any systematic differences across cohorts in each calendar year that

[17] In the time period we examine, no state repeals a duty-to-bargain law.
[18] We choose this event window because the sample sizes become small for relative time indicators less than -10 and greater than 20. Including these "catch-all" relative time indicators allows us to use the full analysis sample, but we caution that it is rather difficult to interpret the coefficients on these two variables.

may be correlated with both the prevalence of DTB laws and labor market outcomes. The state fixed effects control for variation in outcomes that are common across birth cohorts within a state, and the year fixed effects account for national shocks that impact all birth cohorts in the same year. We also control for the proportion of each state-cohort-year-gender cell that is black, Asian, Hispanic or "other." These controls are in the vector *X* in equation (1).

The parameters of interest in equation (1) are π_0-π_{20}, which show the long-run effects of DTB laws among cohorts who are first exposed to these laws in relative years 0 to 20. We show a full set of π estimates in the figures below, but to summarize our findings in a parsimonious way we present effects at 5 (π_5), 10 (π_{10}) and 15 (π_{15}) years in the tables. Effects at 10 years are our preferred estimates because they show the effects among the first cohort that spent nearly the entirety of its schooling years in a DTB environment.

Conditional on the controls in the model, the variation in duty-to-bargain law exposure comes from two sources. The first is within-state differences in exposure over time driven by the state's year of DTB law passage. The second is cross-state variation in the timing of when states passed these laws or whether they pass such a law. The assumptions underlying the identification of parameter π_0-π_{20} are similar to all difference-in-difference analyses: the decision of whether and when to pass a DTB law must be uncorrelated with any prior trends in outcomes across birth cohorts within each state, and the timing of law passage cannot coincide with any state-specific shocks that are isolated to the treated cohorts or with other policies that might influence long-run educational attainment or labor market outcomes.

The π_{-10} to π_{-2} estimates in equation (1) allow us to test the assumption that there is no selection on fixed trends across cohorts. If outcomes are trending in the direction of the estimated treatment effects prior to passage of DTB laws, it suggests a bias from secular trends. As a further check on the credibility of this assumption, we estimate parametric event study models in which the treatment effect is identified relative to a linear pre-treatment trend.

The estimates are very similar to those from equation (1), which indicates little bias from secular trends.

The second potential identification problem of unobserved state-cohort specific shocks correlated with the passage of duty-to-bargain laws is more difficult to investigate. However, there is much variation in the timing of the passage of these laws, as shown in both Figure 1 and Table 1, making it very unlikely that there are secular shocks that are systematically correlated with the timing of DTB passage and only influence the affected cohorts. Permutation tests further support the contention that unobserved shocks correlated with the timing of the rollout of DTB laws are not biasing our estimates. We also include a robustness check that includes state-by-year fixed effects. While less precise, these results indicate that our estimates are not being influenced by state-specific macroeconomic shocks or current statewide policies.

The existence of alternative policies that were passed concurrently with DTB laws is a more serious threat to identification. The 1960s-1980s saw many changes to both schooling and social policies that could have affected the birth cohorts we analyze. If the rollout of these policies is correlated with DTB passage, it could bias our results. We address this concern by controlling for exposure to three alternative policies that occurred concurrently with the DTB movement that also could impact these students' long-run outcomes: school finance reform, the earned income tax credit (EITC), and food stamps. We know of no other policy changes that could plausibly have impacted the declines in labor market outcomes we document. In the vector X in equation (1), we control for the number of years each birth cohort would have been exposed to legislative or court-ordered school finance reform (separately) while in school. The timing of legislative and court-ordered school finance reform are taken from Jackson, Johnson and Persico (2015), who show these reforms led to large increases in the outcomes we consider. We also control for average state EITC rates between the ages of 6 and 18 for each

cohort, as Bastian and Michelmore (forthcoming) show that these policies positively affect educational attainment.[19] Finally, Hoynes, Schanzenbach and Almond (2016) demonstrate that exposure to the food stamp program when young has long-run effects on health and economic outcomes. We use the population-weighted average proportion of counties eligible for food stamps when each birth cohort-state of birth group was between 6 and 18.[20] Below, we show estimates both with and without these controls; they have little effect on our results.

5. Results

Tables 2-4 present baseline estimates of the effect of teacher collective bargaining exposure on labor market outcomes for men (columns i-iii) and women (columns iv-vi) in exposure years 5 (π_5), 10 (π_{10}) and 15 (π_{15}). These estimates show changes in outcomes relative to the year prior to DTB passage, which is set to zero in the event study models. Each column in each panel comes from a separate estimation of equation (1), and we add controls sequentially across columns. In columns (i) and (iv), we control for birth state, birth cohort and calendar year fixed effects as well as race/ethnicity. We add controls for state EITC, school finance reform and food stamp exposure during childhood in columns (ii) and (v), and columns (iii) and (vi) adds cohort-by-year (i.e., age) fixed effects. We discuss the estimates for men and women in turn below.

[19] Cohodes et al. (2016) and Brown, Kowalski and Lurie (2015) show that the Medicaid expansions of the 1980s and 1990s had large, positive effects on the educational attainment and eventual earnings of youth exposed to these expansions. However, our birth cohorts are mostly too old to have been impacted by these policy changes. Furthermore, we cannot control for Medicaid eligibility in this study because eligibility policies and rates are not available prior to 1980. If anything, this is likely to attenuate our estimates.

[20] The food stamp data come from Hoynes, Schanzenbach and Almond (2016), available at https://assets.aeaweb.org/assets/production/articles-attachments/aer/app/10604/20130375_app.pdf.

5.1. Baseline Male Estimates

Table 2 presents results for earnings (Panel A) and hours worked (Panel B), both of which include zeros. Across the first three columns in Panel A, there is clear evidence of a negative effect of teacher collective bargaining on male earnings that grows with exposure time. The estimate in column (iii) indicates that attending school in a state with a duty-to-bargain law for 5 years reduces earnings by $1,542.62 dollars per year. The effect grows to -$1,973.48 in year 10 and $-2,553.31 in year 15. We focus on the effect at year 10 because it represents exposure for nearly all schooling years. The reduction in earnings among the 10-year cohorts is 3.64% relative to the mean, which is shown directly below the estimates in the table. The 3.64% reduction in annual earnings for each individual translates to a large amount of total earnings lost because of the prevalence of duty-to-bargain laws in the US. Across all 33 states that have a duty-to-bargain law in place, our results suggest an *annual* loss of $198.1 billion dollars due to male workers having grown up in states that mandate collective bargaining between teachers' unions and school districts.[21] As the 15 year estimates show, this is likely a conservative estimate of earnings losses due to duty-to-bargain exposure. Furthermore, the estimates in Table 2 are similar across columns, which is inconsistent with biases from age-specific shocks or from exposure to other policies when young.

Panel (a) of Figure 2 shows the full set of event study estimates for male earnings.[22] We overlay a linear fit for the pre- and post-treatment periods to see if there are differential pre-treatment trends and if there are time-varying treatment effects. In Section 5.4, we show estimates that test directly for biases associated with any pre-treatment trends. The visual evidence in Panel (a) of Figure 2 supports our identification strategy: there is no evidence of

[21] We obtain this estimate using total wage income for each state and the percent of the workforce that is male (53.16%) in 2014, obtained from the Bureau of Labor Statistics. Specifically, we multiply 2014 total income in the 33 states by 0.0364*0.5316.

[22] The event study estimates are based on an unbalanced panel of states due to the timing of when our outcomes are measured and the timing of DTB passage. In results available upon request, we have estimated event studies using the small set of states for which we have sufficient pre- and post-DTB observations. The estimates continue to show no signs of pre-treatment trends, and the effect sizes are somewhat larger. There is no evidence that the unbalanced panel we use throughout drives our results and conclusions.

differential trends in earnings across pre-treatment cohorts. When duty-to-bargain laws are passed, earnings decline rather linearly as a function of exposure time. The treatment effect patterns shown in Table 3 from the 5, 10 and 15 year effects thus provide an accurate depiction of how DTB law exposure affects earnings.

Panel B of Table 2 presents estimates for weekly hours worked. Consistent with the reduction in earnings, average hours worked decline by 0.426 due to being exposed to DTB laws for 10 years. This is a 1.09% decline relative to the mean of 38.96 shown in Table C-1. The estimates are stable across columns and are significant at the 5% level for men. As with earnings, the negative effect grows linearly in magnitude with years of exposure from a small and not statistically significant effect at year 5 to -0.655 hours in year 15. Panel (c) of Figure 2 shows event study estimates for this sample and outcome: there is no evidence of differential pre-treatment trends, and the effect grows linearly with relative time.[23]

That teacher collective bargaining is associated with fewer working hours among men suggests that DTB laws may affect the extensive margin of labor supply. Table 3 examines this question, showing estimates of equation (1) where the proportion employed (Panel A), unemployed (Panel B) and not in the labor force (Panel C) are used as dependent variables. DTB laws reduce male employment and increase the proportion of male workers who are not in the labor force. In Panel A, 10 years of exposure to a DTB law lowers the likelihood a male worker is employed by 0.9 of a percentage point, or 1.09% relative to the mean. The estimates are significant at the 5% level and are similar in magnitude to the hours worked results. Thus, much of the reduction in hours worked is coming from the extensive margin.[24]

---

[23] Event study estimates for hours worked in Figure 2 as well as for employment outcomes in Figure 3 show evidence of a shift in the year just prior to DTB passage. As discussed in Section 4, some of this shift is due to misclassification of treatment timing across cohorts because we do not know the year in which respondents entered school. It is unlikely these level shifts represent systematic shocks because of the time-varying nature of the treatment. Importantly, all of our event study estimates reported in the tables are relative to year -1, which is set to zero. Thus, our estimates reflect the change in slope at DTB passage rather than any level shift that occurs prior to passage. Furthermore, the changes between relative years -2 and -1 are not indicative of broader pre-treatment trends in the direction of the treatment effect.

[24] That there is an extensive margin effect makes it difficult to examine wages. We therefore focus on earnings,

There is little evidence of an effect on unemployment. Rather, teacher collective bargaining laws impact labor force participation: 10 years of exposure to a duty-to-bargain law reduces the male labor force participation rate by 0.8 of a percentage point. Relative to the mean labor force non-participation rate, this represents a reduction of 6.56%. As with the results in Table 2, effects at year 15 are even larger than those at year 10.

Full event study estimates of employment outcomes are shown in Figure 3. Pre-treatment trends are small and if anything are in the opposite direction of the treatment effects. As with hours worked in Figure 2, there is a level shift that occurs two years before treatment. But the estimates in Table 3 reflect only the post-DTB trend break. Thus, these estimates are conservative as they do not incorporate the level shift that occurs right before DTB passage. The figure shows clear effects of DTB passage on employment and labor force participation that grows over time, but there is no evidence of an effect on unemployment.

Table 4 presents results for occupational skill and educational attainment. In Panel A, the dependent variable is the proportion of individuals in one's occupation that has at least some collegiate attainment.[25] The results suggest that being exposed to a duty-to-bargain law for 10 years decreases the proportion of workers in one's occupation with at least a college degree by 0.003 (or 0.48% relative to the mean) in our preferred model. While the year 10 effect is not statistically significantly different from zero at conventional levels, both the year 5 and year 15 estimates are of similar magnitude and are significant at the 10% level. Panel (a) of Figure 4 shows full event study estimates for this outcome. The figure shows no evidence of pre-DTB differential trends, and there is a reduction in occupational skill post law passage that accords closely with the estimates in Table 5.[26] These results point to collective

[25] The regressions in Panel A of Table 4 are estimated using the individual-level, disaggregated ACS data. This was done because the dependent variable does not lend itself simply to aggregation at the state-year-cohort level.

[26] Figure 4 shows that much of the effect of occupational sorting occurs immediately, with much smaller growth in the magnitude of the effects over time since DTB passage than we document for other outcomes. Thus, the pattern of effects for this outcome

bargaining laws negatively affecting the occupational skill level chosen by workers.

The reduced earnings and labor force participation associated with teacher collective bargaining suggest that human capital accumulation is declining among exposed cohorts. This reduction could show up in changes in the quantity of education completed, although educational attainment is a coarse measure of human capital. We examine how exposure to a DTB law affects years of completed education; estimates on non-cognitive outcomes that provide alternative measures of human capital are shown in Section 6. Because most people have finished their formal schooling by their mid-30s, the age ranges included in our analysis allow us to measure the total amount of education obtained by each ACS respondent.

Panel B of Table 4 shows results for the total number of years of education. Across columns, the point estimates are negative, modest in magnitude, and are only statistically significantly different from zero at 15 years. Taking the estimates at face value, they suggest a 0.16% decline in educational attainment at 10 years that roughly doubles at 15 years due to collective bargaining exposure.[27] The event study estimates in Panel (c) of Figure 4 indicate a somewhat stronger result. There is a small upward pre-treatment trend that biases the estimates in Table 4 towards zero.

How much of the earnings decline can the educational attainment effects explain? The 10-year estimate is precise enough to rule out an effect larger than -0.120 years of completed schooling at the 5% level in column (iii), which is 0.90% relative to the mean. Assuming that an additional year of schooling increases earnings by 10% (Card 1999), changes in completed

differs from the other labor market outcomes we examine. This likely reflects somewhat different mechanisms driving the occupational sorting results than the other labor market effects.

[27] Examining total years of schooling may miss heterogeneous effects across the distribution of schooling levels. We have estimated equation (1) using the proportion of respondents with different highest levels of educational attainment as the dependent variable to check whether total years of schooling is masking shifts at particular parts of the educational attainment distribution. The estimates for all schooling levels are small in absolute value and only the effect on "some college" is significant at the 5% level. The small negative effect on reduced years of education thus predominantly reflects lower college enrollment, but we cannot rule out small declines that are distributed evenly throughout the educational attainment distribution. These results are available from the authors upon request.

schooling can explain at most 33% (0.12/0.36) of the earnings effect we find.[28] The earnings effect also is likely driven to some extent by changes in teacher quality, both from changes in who become a teacher and teacher effort. Chetty, Friedman, and Rockoff (2014) show that having a one standard deviation higher value-added teacher in one grade increases earnings at age 28 by 1.3%. Assuming that teacher value-added effects are cumulative across grades, our earnings effect is consistent with a 0.28 (3.64/(10*1.3)) reduction in teacher value-added.

The lack of strong educational attainment effects is somewhat surprising, especially given the large labor market effects we document. These results are consistent with some of the prior literature discussed in Section 2 that has not found an effect of duty-to-bargain law passage on high school dropout rates (e.g., Lovenheim 2009). The implication of the educational attainment results is that collective bargaining law exposure affects human capital in ways that are not fully captured by years of education or degree receipt. Our estimates likely reflect other aspects of human capital accumulation that do not appear in educational attainment measures, such as non-cognitive skills, and they highlight the value of examining labor market measures in order to draw a more complete picture of how teacher collective bargaining affects long-run outcomes. We return to this issue in Section 6 when we discuss effects on non-cognitive outcomes.

Our results suggest that male students experience worse long-run labor market outcomes when exposed to duty-to-bargain laws. As discussed previously, we are unable to fully examine the mechanisms that underlie this result due to lack of information on teacher productivity and only sparse data on schooling inputs from this time period. Our results are consistent with Frandsen (2016), who shows that DTB law passage leads to fewer work hours among teachers. Litten (2017) also finds evidence from the restriction of collective bargaining

[28] One concern with the estimates in Table 4 is that the ACS changed the way it asked about the total number of years of schooling in 2008. We estimate equation (1) for the total years of schooling outcome using data only from 2008-2012 in Appendix Table C-3. The estimate is somewhat larger in absolute value but qualitatively similar to the baseline estimate in Table 5. The estimate in Table 5 also is within the 95% confidence interval of the estimate in Table C-3.

rights in Wisconsin that teacher non-wage compensation is reduced. Using the Census/Survey of Governments from 1972-1991, we estimate models of DTB law passage on state average schooling resource allocations that allow for linear pre- and post-DTB trends as well as a level shift in the year of passage (see equation 2). Appendix Table C-9 presents evidence that DTB passage increases the total amount spent on teachers, especially relative to a negative pre-passage trend, but the largest effect is on administrative salary expenditures.[29] These expenditures increase dramatically following law passage, but total expenditures do not change. The shift toward teaching and administrator salaries come at the expense of support service salaries. That the effect grows over time matches the pattern of results in the event study models closely. It is plausible that these changes could reduce school productivity, but we are unaware of research demonstrating a clear link between spending on school administration and student achievement. We also find no effect on teacher-student ratios.

5.2. Baseline Female Estimates

Tables 2-4 and Figures 2-4 also show results for women. The results shown in the tables are suggestive of a small negative effect of collective bargaining law exposure among women on labor market outcomes. Importantly, the event study estimates in Figures 2-4 indicate that these effects are biased by cross-cohort pre-DTB trends that are in the same direction as the treatment effects. Unlike the results for men, the pre-trends among women indicate that any negative effects we find are spurious. We therefore urge caution in lending a causal interpretation to these findings.

The pre-treatment trends among women likely reflect strong secular shifts in female labor market opportunities over the cohorts we consider (Blau and Kahn 2013; Bick and Bruggeman 2014). The shifts happen to be negatively correlated with the timing of DTB

[29] Prior research using these data examine average teacher salaries, not total spending on teachers. This can account for some of the differences between these estimates and those in Hoxby (1996) and Frandsen (2017).

passage, but it is clear that the forces driving these trends do not affect male outcomes; we find no evidence of a bias from such trends for males either visually or statistically when we control for cross-cohort pre-DTB outcome trends in Section 5.4. Thus, the data are inconsistent with an effect of DTB law exposure among women on labor market outcomes, but there is a clear negative effect for men. Motivated by these findings, we focus much of the remainder of the analysis on men but also present female estimates for completeness.

5.3. Estimates by Race/Ethnicity

We show estimates by race and ethnicity at 10 years in Table 5. Panels A and B present results for black and Hispanic men and white and Asian men, respectively, and Panels C and D present similar results for women. Examining results among blacks and Hispanics separately is of great interest, as urban areas that differentially service minority students were more likely to unionize first and to have stronger unions.[30] Furthermore, the 1980s saw a relative erosion of labor market outcomes of young black men (Bound and Freeman 1992). This was a time period in which many of those exposed to a DTB law were entering the labor market, and examining effects for non-whites versus whites could reveal substantial heterogeneity in treatment effects.

As shown in Panel A, the impact of duty-to-bargain law exposure is particularly large among black and Hispanic men: at 10 years earnings decline by $3,019 (8.8%), hours worked decline by 0.58 (1.8%), employment declines by 0.9 of a percentage point (1.3%), and labor force non-participation is reduced by 1.2 percentage points (5.7%). We also find a statistically significant decline in years of schooling of 0.20 years (1.6%) and a significant decline in occupational skill. Panel (a) of Figure 5 presents earnings event study estimates for this

[30] Urban districts were more likely to be represented by the more confrontational AFT rather than the NEA, which could drive some of our results. It also could be that teachers unions themselves have different effects on non-white children. Unions could exacerbate racial differences in disciplinary behavior or otherwise lead to differences in how African American and Hispanic children are treated relative to white and Asian children. Investigating this mechanism is beyond the scope of the paper, but the reductions in non-cognitive skills we show in Section 6 are consistent with this mechanism.

sample. Event studies for other outcomes are presented in Appendix Figures A-1 through A-5. For each outcome, pre-DTB trends are either zero or in the wrong direction (i.e., opposite to the direction of the treatment effect), and the effect grows with more exposure to a collective bargaining law. In short, these figures mirror the event study estimates for the male sample as a whole but are larger in magnitude.

Panel B of Table 5 shows that the estimates are not isolated to black and Hispanic men; statistically significant adverse effects are present for white and Asian men at 10 years as well, though they are more modest in magnitude. Earnings among white and Asian men decline by $1,531 (2.6%) at 10 years of exposure, and employment declines by 0.23 of a percentage point (0.57%). The other estimates are consistent with a decline in outcomes and are similar in magnitude to the baseline estimates.

Comparing Panels (a) and (c) of the race-specific event study figures shows that duty-to-bargain laws lead to worse labor market outcomes among blacks and Hispanics that grow over time, while for whites and Asians the effect is more immediate for several of the outcomes. Hence, the growth in effect sizes with DTB exposure in the baseline estimates is driven predominantly by black and Hispanic men.

Results in Panels C and D of Table 5 show suggestive evidence of DTB exposure on outcomes of black and Hispanic women. However, for several outcomes there are differential pre-treatment trends in the same direction as the treatment effect among these women. These trends are not present for all outcomes, but the results in Panel C of Table 5 should be interpreted with caution given the event study results. That there is evidence of a negative effect of DTB laws among black and Hispanic women indicates that duty-to-bargain laws have large negative impacts on non-whites. The evidence of effect heterogeneity across race/ethnicity for both men and women suggest collective bargaining laws exacerbate long-run racial inequality in outcomes.

5.4. Robustness Checks

The baseline estimates support the rent-seeking theory of union behavior, whereby unions reduce the productivity of public schools and cause a reduction in student achievement as well as subsequent long-run labor market outcomes. In this section, we explore evidence on whether our results are driven by other policies, trends or events that are not accounted for by the controls in equation (1).

We first show results from estimates of parametric event study models that directly control for pre-DTB trends. We construct a relative time to DTB law variable ($C - t_0 + 18$) that forms the basis for the relative time indicator variables in equation (2).[31] This variable takes on a value of zero in states that do not pass a duty-to-bargain law. We then estimate models of the following form:

$$Y_{sct} = \alpha_0 + \alpha_1(C - t_0 + 18)_{sc} + \alpha_2 I(DTB)_{sc} + \alpha_3(C - t_o + 18) * I(DTB)_{sc} + \gamma X_{sct} + \delta_{ct} + \theta_s + \phi_t + \varepsilon_{sct}. \quad (2)$$

All variables are as previously defined. In equation (2), we allow for a level shift (α_2) and a slope shift (α_3) relative to any pre-treatment trend (α_1). Thus, this model is not biased by linear pre-DTB trends, so comparing these estimates to baseline provides some evidence of the importance of directly controlling for cross-cohort variation prior to DTB law passage.

Results of estimating equation (2) are shown in Table 6. The results align with the event study estimates and indicate that our results for men are not biased by pre-treatment trends. For only one outcome is there a significant pre-treatment trend estimate, and it is in the

[31] Similar to the event study estimates, we group relative time observations less than -10 and greater than 20 together. We do so to make this model as similar as possible to equation (1) and to avoid the estimates being unduly influenced by observations that are far away from the timing of treatment. This ensures we are identified off the 30 year period surrounding duty-to-bargain law passage.

opposite direction of the treatment effect. Aside from unemployment and years of education, there are both level and slope shifts that are of similar magnitudes to those in the baseline tables and that mirror the event study plots. We calculate percent effects after 10 years ($(\alpha_2 + \alpha_3 * 10)/\bar{Y}$), which are directly comparable to the percent effects shown in Tables 2-4. These calculations show estimates that are similar to the baseline results.[32]

Panel B shows estimates of equation (2) for women; similar to the event studies there are pre-treatment trends that undermines the validity of the analysis for women. Conditional on these linear trends, there is little evidence of an effect on female labor market outcomes.

Appendix Tables A-4 and A-5 present additional robustness checks that each examines how our results and conclusions for men and women, respectively, change when we control for additional factors in equation (1) that could be correlated with both duty-to-bargain exposure and long-run outcomes. Throughout, we focus on the 10-year estimates; full event study results for each specification are available upon request.

In Panel A, we exclude the 14 states that do not have anti-strike penalties associated with their duty-to-bargain laws.[33] Teacher strikes may have an independent effect on student outcomes, and there is some evidence that resource effects of unions were larger in such states (Paglayan forthcoming). It also could be the case that states becoming more favorable to teachers' unions were becoming more favorable to private sector unions. In Panel B, we control for the total unionization rate at age 18 for each birth state-birth cohort.[34]

The next two panels address the possibility that the rollout of duty-to-bargain laws is correlated with inner-city violence and white flight that occurred during the 1960s and 1970s.

[32] The estimates using equation (2) are somewhat larger because they include to some extent the changes in outcomes that occur between relative periods -2 and -1, which are evident in the event study figures. This illustrates the value of using a less parametric model (such as equation 1) that can better disentangle changes that occur post treatment from those that occur just prior to treatment.

[33] These states are Wisconsin, Connecticut, Michigan, Massachusetts, Rhode Island, Maine, Vermont, Alaska, Hawaii, Kansas, Pennsylvania, Idaho, Oregon and Montana.

[34] Unionization rates come from CPS Merged Outgoing Rotation Group data collected by Barry Hirsch and David Macpherson: http://www.unionstats.com.

Such events likely had independent negative effects on long-run outcomes, which could be driving many of our results. First, we control for the average proportion of people in each state living in urban areas during each cohort's schooling years.[35] While we do not know if a respondent grew up in an inner city, the bias stemming from secular shocks occurring within cities should be correlated with the proportion of individuals living in inner-city areas. Furthermore, this control helps account for increasing suburbanization that was occurring when our analysis cohorts were in school. Second, we use data on all riot and collective action protest events from the Dynamics of Collective Action dataset that includes counts of all collective action events from 1955-1995. We count the number of riots as well as the number of protests in which violence occurred in each state over the time period when each cohort was between 6 and 18.[36] This specification is designed specifically to examine the effect that the urban civil unrest in the 1960s and 1970s has on our estimates. Panel D contains the results that include this additional control. All results in Panels A-D are extremely similar to baseline.

In Panel E, we control for both state-of-birth and current state-of-residence fixed effects (Card and Krueger 1992a,b). The latter set of fixed effects account for the different labor markets in which workers are located that could be correlated with treatment. We estimate this model with individual-level disaggregated data, and the results are larger in absolute value than baseline: not accounting for current state of residence leads to more conservative estimates.

Panel F adds controls for state-by-year fixed effects. These estimates account for any birth state specific shocks or policies that affect all birth cohorts similarly in a state and year. The estimates are noisier than in the baseline models, but they are qualitatively similar and

[35] Urban areas include those living in "urbanized areas" or in "incorporated places"/Census Designated Places (areas with a population of 2,500 or more outside of an urbanized area). This proportion is calculated using the 1960-1990 Decennial Censuses. We use each decennial Census estimate and average across cohorts using the percentage of their school-age years spent in each decade.

[36] This dataset can be found at: http://web.stanford.edu/group/collectiveaction/cgi-bin/drupal/. We obtain similar results if we control for the number of collective action protest events including nonviolent events.

somewhat larger. These results are consistent with our preferred estimates and provide no evidence of bias from state-by-year specific shocks. Panel G complements these findings by showing estimates in which we control for Census Region-by-cohort fixed effects. Some regions may be experiencing differential shocks during the time period in which these laws are passed, such as desegregation in the south. The estimates in Panel G use only within-region and cohort variation, and they are extremely similar to the baseline results if somewhat larger in absolute value. Finally, in Panel H we control for the proportion of time in each cohort's schooling years and state that Democrats had majority control of the state legislature. We do this in order to account for the potential correlation between political control of the state legislature and unionization. The similarity of the estimates suggests we are not picking up political trends of shifts that drive long-run labor market outcomes.

We also examine the sensitivity of our results to outliers by re-estimating equation (1) 50 times for all of our outcomes, each time dropping a different state from the analysis sample. The results from this exercise are shown in Appendix Figure C-6 for four of our main outcomes: earnings, hours of work, employment, and labor force participation. Our male estimates are insensitive to excluding any one state.[37]

As discussed above, of primary concern in our identification strategy is the existence of secular trends that differ systematically with treatment exposure. The event study estimates for men suggest such trends are not biasing our estimates. As an additional test of whether the timing pattern of DTB passage is driving our results, we perform permutation tests for all of our outcomes that randomly reassign DTB passage years across states. We do this in two ways: first, we randomly assign DTB passage dates between 1960 and 1987 across states, and

[37] Because of the geographic concentration of DTB rollout, we lack the power to estimate models separately by Census region or that drop specific regions. We also lack the power to drop states that never pass DTB laws (many of which are in the south). The estimates in Panel G of Table C-4 indicates that our results are not driven by region-specific trends or shocks or by the inclusion of any specific region in our sample, and the estimates in Figure C-6 suggest our results are not being driven by the inclusion of any one state.

second we randomly assign DTB passage dates to match the timing distribution shown in Figure 1. Table 7 shows the results from this exercise for men. We perform the permutation test 300 times for each outcome and calculate the percentage of times the simulated estimate is less than the actual estimate. These results therefore represent p-values of the null hypothesis that any combination of passage dates across states would generate the same pattern of treatment effects. We reject such a null at either the 5 or 10 percent level for every outcome except years of education in Panel A. These results suggest that our baseline estimates are not identified off of secular trends or endogenous timing of DTB passage. That the effects we estimate are linked strongly to both whether a state passes a DTB law and when it does so supports the validity of our estimation strategy.

A final identification issue comes from measurement error driven by either pre- or post-birth mobility. To assess the importance of pre-birth mobility, we estimate equation (1) using observed fixed characteristics in the ACS and some state-year level observables that are unlikely to be affected by teacher collective bargaining. Because we focus on state of birth, these estimates show whether the composition of people born in a given state and cohort changed with respect to duty-to-bargain law exposure. Table C-6 shows these results. We find little evidence of a change in the composition of birth cohorts that would indicate parents are systematically moving prior to having a child because of duty-to-bargain laws. While there is a small number of statistically significant coefficients, they are quite close to zero and thus economically insignificant. The point estimates are also not in a consistent direction that would indicate a bias from changes in composition driven by DTB law passage.

We next examine the relevance of post-birth mobility, which introduces measurement error into our DTB exposure variable. In the 1990 Census, 78.4% of 17-year-olds live in the state of their birth. In order to provide information about how serious any mobility-induced bias would be, we re-estimate equation (1) under two assumptions. In Panel A of Appendix

Table C-7, we show results for men that exclude the 37.7% of respondents who do not live in their birth state.[38] These estimates are typically larger in absolute value than our baseline estimates, although they are close in magnitude.

In Panel B, we estimate equation (1) under the assumption that those who live in a state at age 17 other than their birth state spent all of their schooling years in that other state. Using the 1990 Census, we create a 50x50 matrix that contains the full joint distribution of state-of-birth and state at age 17. We then create a new dataset that contains 50 observations for each age-year-birth-state observation. Within each age-year-birth-state group, there is a separate observation for each potential state a respondent could have lived in at age 17. We then weight each observation by the proportion of the 1990 Census that was in the given birth state-state at 17 combination. All DTB and other state-specific variables are calculated using the assumed state at age 17, not the birth state. Standard errors are two-way clustered at the birth state, state at age 17 level (Cameron, Gelbach and Miller 2011).[39] The results in Panel B are very similar to baseline in magnitude and statistical significance. Taken together, the results in Table C-7 suggest that any bias from post-birth mobility is small.

6. Medium-Term Effects on Non-Cognitive Outcomes

The negative effects of teacher collective bargaining on earnings and labor force participation suggest that duty-to-bargain laws lead students to obtain less human capital when in school. We now turn to direct evidence on how collective bargaining influences non-cognitive outcomes using data from the NLSY79. This is a nationally-representative dataset of students aged 14-22 in 1979, covering the 1957-1965 birth cohorts. These cohorts thus overlap with much of the variation in the passage of teacher collective bargaining laws shown in Figure 1.

Non-cognitive skills are measured three ways: the Rotter Locus of Control, the

[38] Estimates for women are shown in Appendix Table C-8.
[39] Because this method requires aggregated data, we do not estimate this model for occupational skill.

Rosenberg Self-esteem Scale and the Pearlin Mastery Scale. The Rotter Locus of Control measures the extent to which students believe they have control over their own lives, with higher scores indicating *less* internal control (i.e., lower non-cognitive skills). The Rosenberg Self-esteem Scale is designed to measure a student's self-worth; higher scores indicate higher self-esteem. Third, the Pearlin Mastery Scale is a measure of the extent to which individuals perceive themselves in control of forces that significantly impact their lives. Respondents with higher measures report increased ability to determine the course of their own life.

We estimate models using these outcomes that are similar to equation (2) but that omit the pre-DTB relative time control due to a lack of a sufficient number of observations. We restrict our analysis to men because of prior evidence of lack of pre-treatment trends and because it is among men where we observe negative labor market effects of DTB law exposure. All outcomes are measured in 1997, so we can only include birth cohort and state of residence at age 14 fixed effects (not birth cohort-year fixed effects). We also control for race, family income and both mother's and father's educational attainment. Estimates are weighted by the NLSY79 sample weights and standard errors are clustered at the state level.

We see consistent evidence in Table 8 that exposure to a collective bargaining law negatively impacts non-cognitive scores among men. All non-cognitive skill measures move in the direction of declining skill: after 10 years, the Rotter Locus of Control increases by 12.9%, the Rosenberg Self-esteem Scale declines by 4.6%, and the Pearlin Mastery Scale score is reduced by 1.6%. The years of exposure estimates are statistically different from zero at the 5% level for the first two measures, while Pearlin Mastery Scale estimates are not significant at even the 10% level. These results show that students exposed to DTB laws experience reductions in non-cognitive skills in adolescence and early adulthood.

The results in Table 8 support the earnings and labor market results presented above. These cognitive and non-cognitive measures have been shown in prior research to be highly

correlated with long-run outcomes (Heckman, Stixrud and Urzua 2006), and they provide more direct evidence consistent with the rent-seeking hypothesis. Teacher collective bargaining laws lead to a decline in the productivity of educational inputs, which reduces short-run non-cognitive outcomes that are still evident into adulthood. Furthermore, these results help explain why the labor market effects of teacher collective bargaining are larger than the educational attainment effects: non-cognitive skills affect the former more than the latter (Heckman, Stixrud and Urzua 2006). The sum total of the evidence from the ACS and NLSY79 is remarkably consistent in showing that teacher DTB laws negatively impact male long-run outcomes through their effects on the quality of education students receive.

7. Conclusion

This paper provides the first analysis of the effect of state teacher DTB laws on student long-run educational attainment and labor market outcomes. We link adults from the 2005-2012 ACS to their state of birth and exploit the timing of passage of duty-to-bargain laws across cohorts within a state and across states over time. Our estimates show that exposure to duty-to-bargain laws when 35-49 year old men were of school-age adversely affects their long-run outcomes. We do not find robust evidence of impacts on women, however.

Our results are consistent with the rent-seeking model of teachers' unions. Men in cohorts who were exposed to a duty-to-bargain law in the 10 years after passage earn $1,973.48 (or 3.64%) less per year. A back-of-the-envelope calculation indicates these laws reduce total labor market earnings by $198.1 billion per year, which suggests our findings have large implications for earnings in the US due to the prevalence of duty-to-bargain laws. Our results also point to collective bargaining laws leading to fewer hours worked, lower employment rates and lower labor force participation. The negative effects of exposure to duty-to-bargain laws are largest among black and Hispanic men, although white and Asian

men also are adversely impacted. In particular, yearly earnings decline by 8.77% among black and Hispanic males. We find more evidence of a decline in educational attainment for this group of men as well. Among white and Asian men, earnings decline by 2.58%. We complement these results with an analysis from the NLSY79 that shows duty-to-bargain laws reduce non-cognitive outcomes among young men. In total, our estimates indicate that state duty-to-bargain laws have sizable, negative labor market consequences for men who attended grade school in states with these laws.

From a policy perspective, these results contribute to the contentious debate occurring in many states about whether to limit the collective bargaining power of teachers. For example, in 2011 Wisconsin, Indiana, Tennessee and Idaho passed legislation that greatly reduced the ability of teachers to bargain with school districts, and in 2014 Michigan passed a public employee right-to-work law that sought to limit union negotiating power. Of first-order concern in this policy debate is how collective bargaining affects student outcomes. Our results provide the most comprehensive information to date on this question. However, there are a couple of caveats to generalizing these findings to current students. First, the cohorts we analyze were exposed to an educational environment very different from the one that exists today. For example, school choice as well as teacher, school and student accountability policies that are currently rather ubiquitous were virtually nonexistent during the 1960s-1980s. Some of the effects of teacher collective bargaining we estimate could be driven by how teachers' unions interacted with specific aspects of the educational system that no longer are relevant. Second, the current collective bargaining law changes in many states alter aspects of collective bargaining, not the legality of collective bargaining itself. Examination of these policy changes will lend much insight into whether one can change collective bargaining laws to reduce the negative impacts on students we find while still providing teachers with the bargaining benefits they value. This is an important set of questions for future research.

References

Balfour, Alan G. 1974. "More Evidence that Unions do not Achieve Higher Salaries for Teachers." *Journal of Collective Negotiations* 3(4): 289-303.

Baron, Jason E. 2017. "The Effect of Teachers' Unions on Student Achievement: Evidence from Wisconsin's Act 10." Mimeo.

Bastian, Jacob and Katherine Michelmore. Forthcoming. "The Long-Term Impact of the Earned Income Tax Credit on Children's Education and Employment Outcomes." *Journal of Labor Economics*.

Baugh, William H. and Joe A. Stone. 1982. "Teachers, Unions, and Wages in the 1970's: Unionism Now Pays." *Industrial and Labor Relations Review* 35(3): 368-376.

Bedard, Kelly and Elizabeth Dhuey. 2012. "School-Entry Policies and Skill Accumulation Across Directly and Indirectly Affected Individuals." *Journal of Human Resources* 47(3): 643-683.

Biasi, Barbara. 2017. "Unions, Salaries, and the Market for Teachers: Evidence from Wisconsin." Mimeo.

Bick, Alexander and Bettina Bruggeman. 2014. "Labor Supply Along the Extensive and Intensive Margin: Cross-Country Facts and Time Trends by Gender." Mimeo.

Blau, Francine and Lawrence Kahn. 2013. "Female Labor Supply: Why is the US Falling Behind?" *American Economic Review* 103(3):251-256.

Bound, John and Richard B. Freeman. 1992. "What went Wrong? The Erosion of Relative Earnings and Employment Among Young Black Men in the 1980s." *Quarterly Journal of Economics* 107(1): 201-232.

Cameron, Colin A., Jonah B. Gelbach and Douglas L. Miller. 2011. "Robust Inference With Multiway Clustering." *Journal of Business and Economic Statistics* 29(2): 238-249.

Card, David. 1999. "The Causal Effect of Education on Earnings." In Orley Ashenfelter and David Card, editors, *Handbook of Labor Economics Volume 3A*. Amsterdam: Elsevier.

Card, David and Alan B. Krueger. 1992a. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy* 100(1): 1-40.

Card, David and Alan B. Krueger. 1992b. "School Quality and Black-White Relative Earnings: A Direct Assessment." *Quarterly Journal of Economics* 107(1): 151-200.

Chetty, Raj, John Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star." *Quarterly Journal of Economics* 126(4): 1593-1660.

Chetty, Raj, John Friedman and Jonah Rockoff. 2014. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104(9): 2633-2679.

Cohodes, Sarah, Daniel Grossman, Samuel Kleiner and Michael F. Lovenheim. 2016. "The Effect of Child Health Insurance Access on Schooling: Evidence from Public Insurance Expansions." *Journal of Human Resources* 51(3): 727-759.

Cowen, Joshua M. and Katharine O. Strunk. 2015. "The Impact of Teachers' Unions on Educational Outcomes: What we Know and what we Need to Learn." *Economics of Education Review* 48(October): 208-223.

Deming, David J., Sarah Cohodes, Jennifer Jennings, and Christopher Jencks. 2013. "School Accountability, Postsecondary Attainment and Earnings." NBER Working Paper No. 19444.

Eberts, Randall W. and Joe A. Stone. 1986. "Teacher Unions and the Cost of Public Education." *Economic Inquiry* 24(4): 631-643.

Eberts, Randall W. and Joe A. Stone. 1987. "Teacher Unions and the Productivity of Public Schools." *Industrial and Labor Relations Review* 40(3): 354-363.

Farber, Henry S. 2003. "Nonunion Wage Rates and the Threat of Unionization." Working Paper no. 472 (March), Industrial Relations Section, Princeton University.

Frandsen, Brigham. 2016. "The Effects of Collective Bargaining Rights on Public Employee Compensation: Evidence from Teachers, Fire Fighters, and Police." *Industrial and Labor Relations Review* 69(1): 84-112.

Freeman, Richard. 1980. The Exit-Voice Tradeoff in the Labor Market: Unionism, Job Tenure, Quits, and Separations." *Quarterly Journal of Economics* 94(4): 643-673.

Goldin, Claudia, Lawrence F. Katz and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap." *Journal of Economic Perspectives* 20(4): 133-156.

Gunderson, Morley. 2005. "Two Faces of Union Voice in the Public Sector." *Journal of Labor Research* 26(3): 393-413.

Haider, Steven and Gary Solon. 2006. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96(4): 1308-1320.

Heckman, James J. and Tim Krautz. 2012. "Hard Evidence on Soft Skills." *Labour Economics* 19(4): 451-464.

Heckman, James J., Rodrigo Pinto and Peter Savelyev. 2013. "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103(6): 2052-2086.

Heckman, James J., Jora Stixrud and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24(3): 411-482.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-run Impacts of Childhood Access to the Safety Net." *American Economic Review* 106(4): 903-934.

Hoxby, Caroline Minter, 1996. "How Teachers' Unions Affect Education Production." *The Quarterly Journal of Economics* 111(3): 671-718.

Hoxby, Caroline Minter and Andrew Leigh, 2004. "Pulled Away or Pushed out? Explaining the Decline of Teacher Aptitude in the United States." *American Economic Review* 94(2): 236-240.

Jackson, C. Kirabo, Rucker Johnson and Claudia Persico. 2015. "The Effect of School Finance Reforms on the Distribution of Spending, Academic Achievement and Adult Outcomes." *Quarterly Journal of Economics* 131(1): 157-218.

Kleiner, Morris and Daniel Petree. 1988. "Unionism and Licensing of Public School Teachers: Impact on Wages and Educational Output," in *When Public Sector Workers Unionize*, Richard Freeman and Casey Ichniowski, eds. (Chicago, IL: University of Chicago Press).

Krueger, Alan. 1999. "Experimental Estimates of Education Production Functions." *Quarterly Journal of Economics* 114(2): 497-532.

Litten, Andrew. 2017. "The Effects of Public Unions on Compensation: Evidence From Wisconsin." Mimeo.

Lott, Jonathan and Lawrence W. Kenny. 2013. "State Teacher Union Strength and Student Achievement." *Economics of Education Review* 35: 93-103.

Lovenheim, Michael F. 2009. "The Effect of Teachers' Unions on Education Production: Evidence from Union Election Certifications in Three Midwestern States." *Journal of Labor Economics* 27(4): 525-587.

Ludwig, Jens and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Quarterly Journal of Economics* 122(1): 159-208.

Moe, Terry M. 2009. "Collective Bargaining and the Performance of Public Schools." *American Journal of Political Science* 53(1): 156-174.

Moore, William J. and John Raisian. 1987. "Union-Nonunion Wage Differentials in the Public Administration, Educational, and Private Sectors: 1970-1983." *The Review of Economics and Statistics* 69(4): 608-616.

Paglayan, Agustina. Forthcoming. "Public-Sector Unions and The Size of Governments." *American Journal of Political Science*.

Quinby, Laura D. 2017. "De-Unionization and the Labor Market for Teachers: From School Boards to State Politics." Mimeo

Roth, Jonathan. 2017. "Union Reform and Teacher Turnover: Evidence from Wisconsin's Act 10." Mimeo.

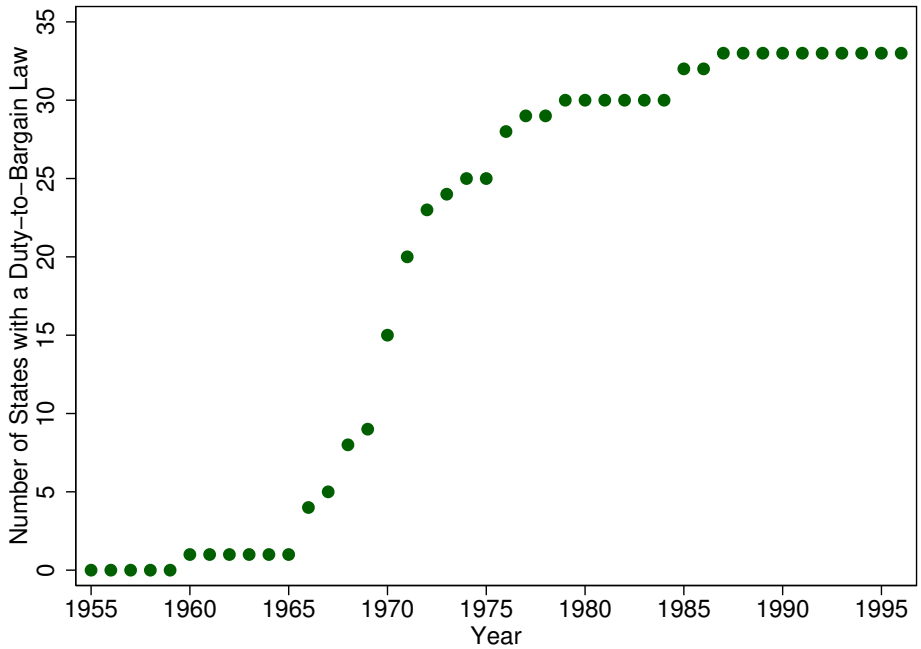
Saltzman, Gregory M., 1985. "Bargaining Laws as a Cause and Consequence of the Growth of Teacher Unionism." *Industrial and Labor Relations Review* 38(3): 335-351.

Valletta, Robert G. and Richard B. Freeman, 1988. "The NBER Public Sector Collective Bargaining Law Data Set." Appendix B in Richard B. Freeman and Casey Ichniowski, (eds.), *When Public Employees Unionize*. Chicago: NBER and University of Chicago Press.

West, Kristine. 2015. "Teachers' Unions, Compensation and Tenure." *Industrial Relations* 54(2): 294-320.

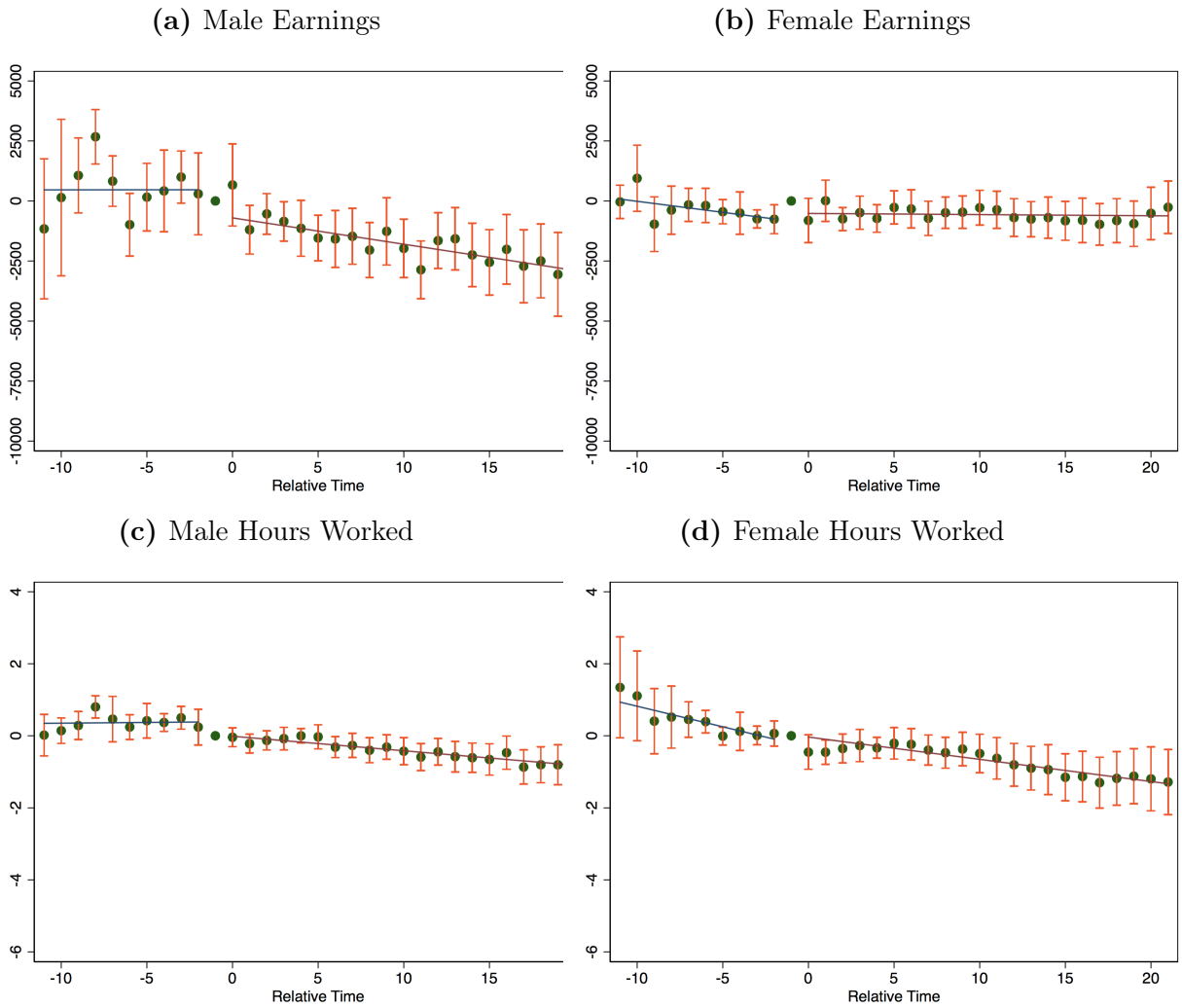
Zuelke, Dennis C. and Lloyd E. Frohreich. 1977. "The Impact of Comprehensive Collective Negotiations on Teachers' Salaries: Some evidence from Wisconsin." *Journal of Collective Negotiations* 6(1): 81-88.

Figure 3.1: The Number of States with Teacher Duty-to-Bargain Laws over Time



Source: NBER Public Sector Collective Bargaining Law Data Set (Valletta and Freeman 1988), updated by Kim Reuben to 1996.

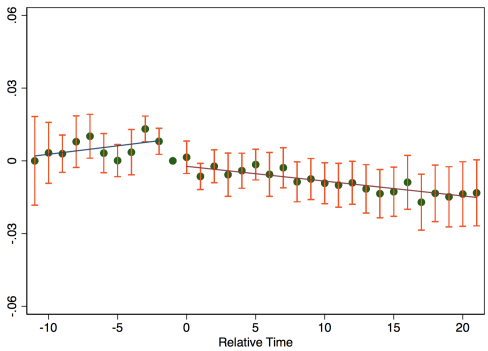
Figure 3.2: Event Study Estimates - Earnings and Hours Worked



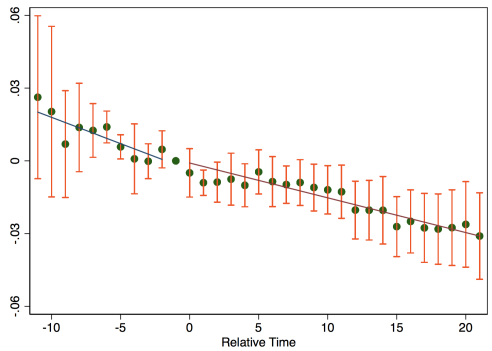
Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Relative year -1 is omitted, so all estimates are in relationship to this year. Relative year -11 includes all observations with relative time $\leq$-11 and relative year 21 includes all observations with relative time $\geq$21. All estimates include birth cohort-by-year, birth state, and year fixed effects as well as controls for the racial/ethnic composition of the state-cohort-year-gender cell and exposure to school finance reforms, state EITC rates, and food stamps. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the state level.

Figure 3.3: Event Study Estimates - Employment Outcomes

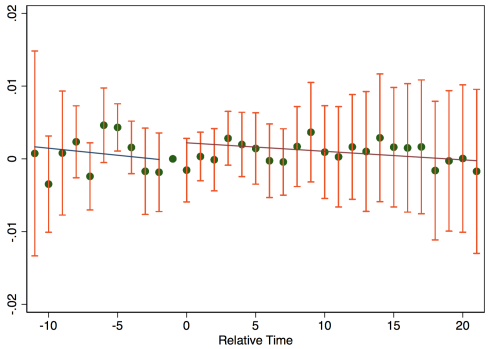
(a) Male Employment



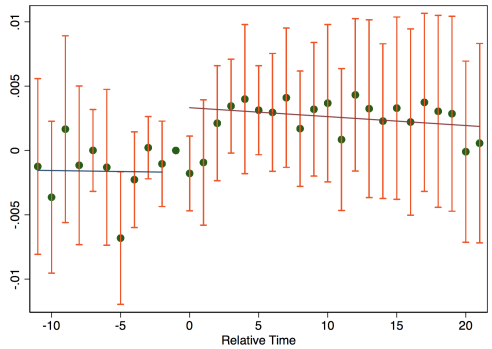
(b) Female Employment



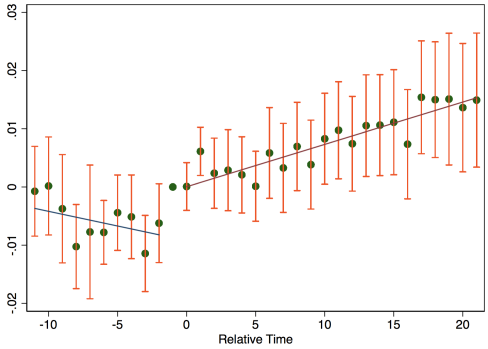
(c) Male Unemployment



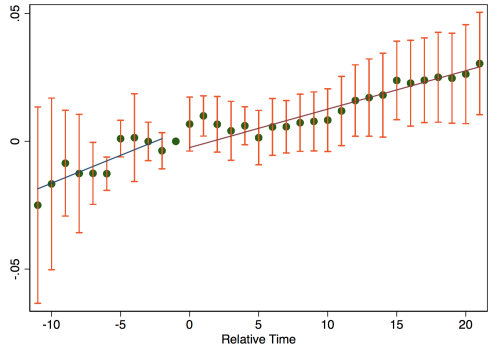
(d) Female Unemployment



(e) Male Not in Labor Force



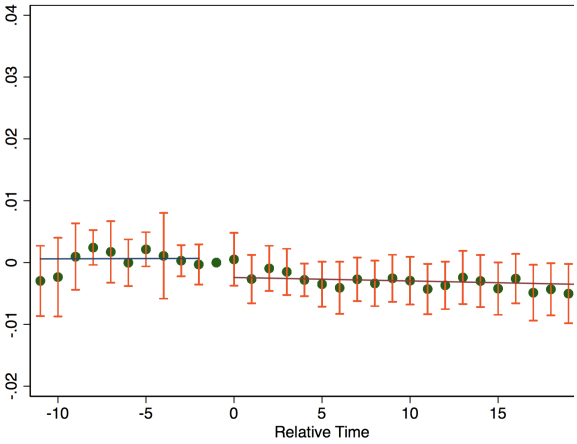
(f) Female Not in Labor Force



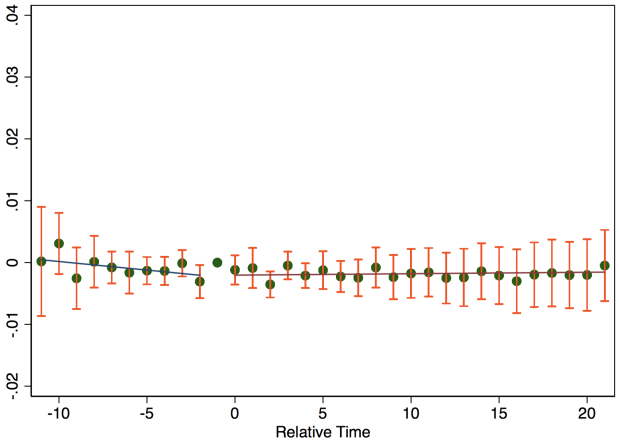
Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Relative year -1 is omitted, so all estimates are in relationship to this year. Relative year -11 includes all observations with relative time $\leq$-11 and relative year 21 includes all observations with relative time $\geq$21. All estimates include birth cohort-by-year, birth state, and year fixed effects as well as controls for the racial/ethnic composition of the state-cohort-year-gender cell and exposure to school finance reforms, state EITC rates, and food stamps. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the state level.

Figure 3.4: Event Study Estimates - Occupational Skill and Years of Education

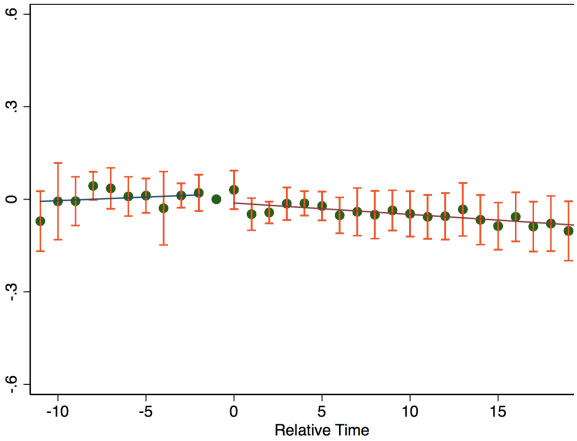
(a) Male Occupational Skill



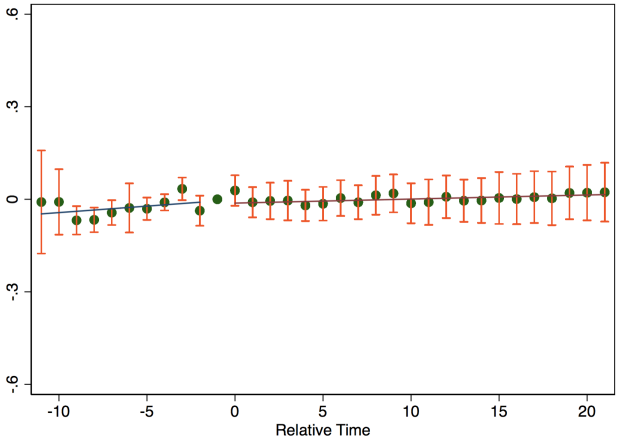
(b) Female Occupational Skill



(c) Male Years of Education

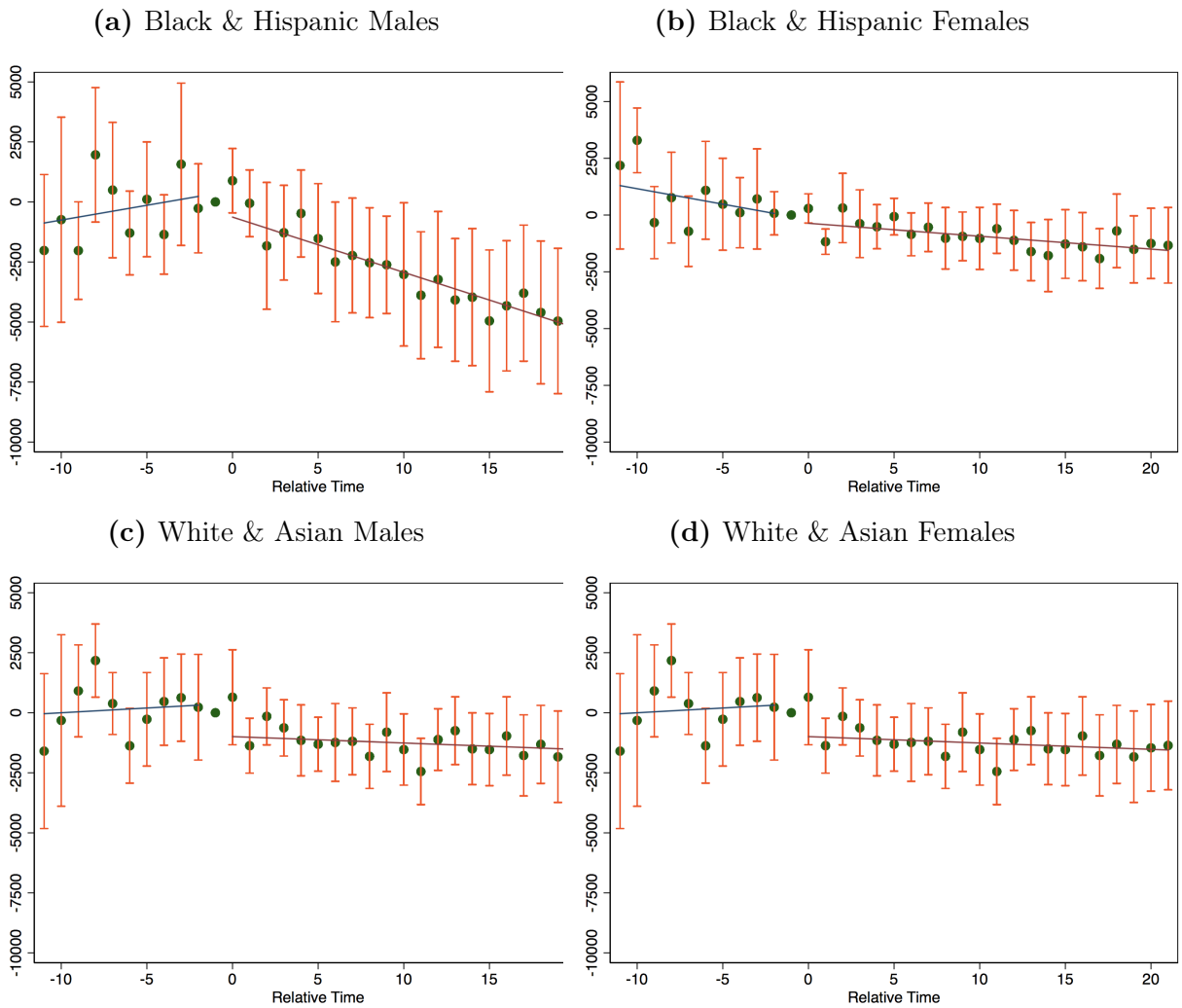


(d) Female Years of Education



Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Relative year -1 is omitted; all estimates are in relationship to this year. Relative year -11 includes observations with relative time $\leq$-11 and relative year 21 includes observations with relative time $\geq$21. All estimates include birth cohort-by-year, birth state, and year fixed effects as well as Other controls for school finance reform, EITC and food stamp measures. In Panels (c) and (d), regressions are based on 6,000 birth state-birth cohort-year observations and include controls for racial/ethnic composition of the state-cohort-year-gender cell. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. In Panels (a) and (b), estimation of equation (1) is done using disaggregated data and includes controls for respondent race/ethnicity. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors clustered at the state level.

Figure 3.5: Event Study Estimates by Gender and Race/Ethnicity - Earnings



Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Relative year -1 is omitted, so all estimates are in relationship to this year. Relative year -11 includes all observations with relative time $\leq$-11 and relative year 21 includes all observations with relative time $\geq$21. All estimates include birth cohort-by-year, birth state, and year fixed effects as well as controls for exposure to school finance reforms, state EITC rates, and food stamps. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-race-gender cell. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the state level.

Table 3.1: Teacher Duty-to-Bargain Law Passage by State

State	Year of Passage	State	Year of Passage
Alabama		Montana	1972
Alaska	1971	Nebraska	1987
Arizona		Nevada	1970
Arkansas		New Hampshire	1976
California	1977	New Jersey	1969
Colorado		New Mexico	
Connecticut	1966	New York	1968
Delaware	1970	North Carolina	
Florida	1976	North Dakota	1970
Georgia		Ohio	1985
Hawaii	1971	Oklahoma	1972
Idaho	1972	Oregon	1970
Illinois	1985	Pennsylvania	1971
Indiana	1974	Rhode Island	1967
Iowa	1976	South Carolina	
Kansas	1971	South Dakota	1971
Kentucky		Tennessee	1979
Louisiana		Texas	
Maine	1970	Utah	
Maryland	1970	Vermont	1968
Massachusetts	1966	Virginia	
Michigan	1966	Washington	1968
Minnesota	1973	West Virginia	
Mississippi		Wisconsin	1960
Missouri		Wyoming	

Source: NBER Public Sector Collective Bargaining Law Data Set (Valletta and Freeman 1988), updated by Kim Reuben to 1996. Blank entries reflect the absence of a teacher duty-to-bargain law in the state.

Table 3.2: The Effect of Collective Bargaining Laws on Earnings and Hours Worked

			Panel A: Earnings			
		Men			Women	
Exposure Time	(i)	(ii)	(iii)	(iv)	(v)	(vi)
At 5 Years	-1738.43***	-1611.59***	-1542.62***	-348.99	-287.64	-269.95
	(476.59)	(483.61)	(484.43)	(355.43)	(352.97)	(352.78)
At 10 Years	-2145.54***	-2042.22***	-1973.48***	-357.26	-295.82	-281.52
	(597.60)	(614.58)	(621.15)	(376.29)	(368.76)	(369.64)
At 15 Years	-2665.56***	-2616.30***	-2553.31***	-899.26**	-838.26**	-824.83**
	(699.07)	(691.33)	(693.83)	(412.58)	(410.47)	(413.05)
% Effect						
At 10 Years	-3.95%	-3.76%	-3.64%	-1.18%	-0.98%	-0.93%
			Panel B: Hours Worked			
		Men			Women	
Exposure Time	(i)	(ii)	(iii)	(iv)	(v)	(vi)
At 5 Years	-0.065	-0.044	-0.029	-0.238	-0.220	-0.210
	(0.164)	(0.173)	(0.170)	(0.220)	(0.222)	(0.222)
At 10 Years	-0.459**	-0.442**	-0.426**	-0.514*	-0.499*	-0.492*
	(0.183)	(0.193)	(0.192)	(0.267)	(0.270)	(0.273)
At 15 Years	-0.676***	-0.673***	-0.655***	-1.148***	-1.153***	-1.151***
	(0.222)	(0.225)	(0.222)	(0.324)	(0.331)	(0.335)
% Effect						
At 10 Years	-1.18%	-1.13%	-1.09%	-1.74%	-1.69%	-1.66%
Other Policy Controls		x	x		x	x
Birth Cohort*Survey Year FE			x			x

Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. 5, 10 and 15 year estimates from the full event study model are shown. Regressions are based on 6,000 birth state-cohort-year observations. All estimates include birth state, birth cohort and year fixed effects, as well as controls for racial/ethnic composition of the state-cohort-year-gender cell. Other Policy Controls include school finance reform, EITC and food stamp measures as described in the text. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. % Effect at 10 Years shows the 10-year effect relative to the mean presented in Table A-1. Standard errors clustered at the birth state level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 3.3: The Effect of Collective Bargaining Laws on Labor Market Participation

			Panel A: Employed			
		Men			Women	
Exposure Time	(i)	(ii)	(iii)	(iv)	(v)	(vi)
At 5 Years	-0.003	-0.002	-0.002	-0.005	-0.005	-0.005
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)
At 10 Years	-0.011**	-0.009**	-0.009**	-0.013**	-0.012**	-0.012**
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)
At 15 Years	-0.014***	-0.013***	-0.013**	-0.027***	-0.027***	-0.027***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
% Effect At 10 Years	-1.34%	-1.09%	-1.09%	-1.78%	-1.64%	-1.64%
			Panel B: Unemployed			
		Men			Women	
Exposure Time	(i)	(ii)	(iii)	(iv)	(v)	(vi)
At 5 Years	0.002	0.001	0.001	0.004**	0.003*	0.003*
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
At 10 Years	0.002	0.001	0.001	0.004	0.004	0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
At 15 Years	0.002	0.002	0.002	0.004	0.003	0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
% Effect At 10 Years	3.51%	3.51%	3.51%	6.90%	6.90%	6.90%
			Panel C: Not In Labor Force			
		Men			Women	
Exposure Time	(i)	(ii)	(iii)	(iv)	(v)	(vi)
At 5 Years	0.001	0.000	0.000	0.002	0.002	0.001
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)
At 10 Years	0.009**	0.009**	0.008**	0.008	0.009	0.008
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)
At 15 Years	0.012***	0.012***	0.011***	0.024***	0.024***	0.024***
	(0.005)	(0.005)	(0.005)	(0.008)	(0.008)	(0.008)
% Effect At 10 Years	7.38%	7.38%	6.56%	3.60%	4.05%	3.60%
Other Policy Controls		x	x		x	x
Birth Cohort*Survey Year FE			x			x

Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. 5, 10 and 15 year estimates from the full event study model are shown. Regressions are based on 6,000 birth state-cohort-year observations. All estimates include birth state, birth cohort and year fixed effects, as well as controls for racial/ethnic composition of the state-cohort-year-gender cell. Other Policy Controls include school finance reform, EITC and food stamp measures as described in the text. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. % Effect at 10 Years shows the 10-year effect relative to the mean presented in Table A-1. Standard errors clustered at the birth state level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 3.4: The Effect of Collective Bargaining Laws on Occupational Skill and Educational Attainment

	Panel A: Occupational Skill					
		Men			Women	
Exposure Time	(i)	(ii)	(iii)	(iv)	(v)	(vi)
At 5 Years	-0.003*	-0.004*	-0.003*	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
At 10 Years	-0.003	-0.003	-0.003	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
At 15 Years	-0.004*	-0.004*	-0.004*	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
% Effect At 10 Years	-0.48%	-0.48%	-0.48%	-0.36%	-0.36%	-0.36%
	Panel B: Years of Education					
		Men			Women	
Exposure Time	(i)	(ii)	(iii)	(iv)	(v)	(vi)
At 5 Years	-0.029	-0.023	-0.022	-0.020	-0.015	-0.015
	(0.025)	(0.024)	(0.024)	(0.027)	(0.028)	(0.028)
At 10 Years	-0.054	-0.047	-0.047	-0.019	-0.013	-0.013
	(0.039)	(0.037)	(0.037)	(0.033)	(0.033)	(0.033)
At 15 Years	-0.091**	-0.087**	-0.087**	0.001	0.004	0.004
	(0.038)	(0.038)	(0.039)	(0.043)	(0.043)	(0.043)
% Effect At 10 Years	-0.22%	-0.17%	-0.16%	-0.15%	-0.11%	-0.11%
Other Policy Controls		x	x		x	x
Birth Cohort*Survey Year FE			x			x

Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. 5, 10 and 15 year estimates from the full event study model are shown. In Panel B, regressions are based on 6,000 birth state-birth cohort-year observations and include birth state, birth cohort and year fixed effects, as well as controls for racial/ethnic composition of the state-cohort-year-gender cell. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. Other Policy Controls include school finance reform, EITC and food stamp measures as described in the text. In Panel A, the dependent variable is the percentage of those in each respondent's occupation with more than a high school degree. Estimation of equation (1) is done using disaggregated data in Panel A and includes birth state, birth cohort and year fixed effects as well as controls for respondent race/ethnicity. % Effect at 10 Years shows the 10-year effect relative to the mean presented in Table A-1. Standard errors clustered at the birth state level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 3.5: The Effect of Collective Bargaining Laws 10 Years Post DTB Passage on Long-Run Outcomes, by Race/Ethnicity

Panel A: Black and Hispanic Men							
Exposure Time	Earnings (i)	Hours Worked (ii)	Employed (iv)	Un-Employed (v)	Not in Labor Force (v)	Years of Education (vi)	Occup. Skill (vii)
At 10 Years	-3019.01**	-0.584	-0.009	-0.003	0.012*	-0.202***	-0.009*
	(1521.46)	(0.435)	(0.010)	(0.007)	(0.007)	(0.060)	(0.005)
% Effect	-8.77%	-1.76%	-1.28%	-3.53%	5.69%	-1.60%	-1.38%
Panel B: White and Asian Men							
Exposure Time	Earnings (i)	Hours Worked (ii)	Employed (iv)	Un-Employed (v)	Not in Labor Force (v)	Years of Education (vi)	Occup. Skill (vii)
At 10 Years	-1530.81**	-0.228	-0.007	0.002	0.005	-0.034	-0.002
	(757.54)	(0.224)	(0.004)	(0.003)	(0.005)	(0.041)	(0.002)
% Effect	-2.58%	-0.57%	-0.82%	4.08%	5.00%	-0.25%	-0.33%
Panel C: Black and Hispanic Women							
Exposure Time	Earnings (i)	Hours Worked (ii)	Employed (iv)	Un-Employed (v)	Not in Labor Force (v)	Years of Education (vi)	Occup. Skill (vii)
At 10 Years	-1026.12	-1.176**	-0.034**	0.020***	0.014	-0.075	-0.007**
	(696.76)	(0.540)	(0.008)	(0.007)	(0.012)	(0.047)	(0.003)
% Effect	-3.92%	-3.92%	-4.83%	27.03%	6.31%	-0.58%	-1.20%
Panel D: White and Asian Women							
Exposure Time	Earnings (i)	Hours Worked (ii)	Employed (iv)	Un-Employed (v)	Not in Labor Force (v)	Years of Education (vi)	Occup. Skill (vii)
At 10 Years	37.11	0.003	-0.001	0.001	0.001	-0.015	-0.001
	(457.76)	(0.282)	(0.007)	(0.002)	(0.008)	(0.044)	(0.002)
% Effect	0.12%	0.01%	-0.14%	2.44%	0.45%	-0.11%	-0.18%

Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. 5, 10 and 15 year estimates from the full event study model are shown. Regressions are based on 6,000 birth state-cohort-year observations. All estimates include birth state, year and birth cohort-by-year fixed effects as well as controls for exposure to school finance reform, food stamps and EITC when of school age. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-birth cohort-year-gender-race cell. % Effects show effects relative to the mean of each dependent variable. Standard errors clustered at the birth state level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 3.6: Parametric Event Study Estimates of the Effect of Collective Bargaining Laws on Long-Run Outcomes

	Earnings (i)	Hours Worked (ii)	Employed (iv)	Un-Employed (v)	Not in Labor Force (v)	Years of Education (vi)	Occup. Skill (vii)
Panel A: Men							
Relative Years to DTB Law	15.19	0.007	0.0001	-0.00004	-0.0001	-0.004	0.0005***
	(67.96)	(0.024)	(0.0005)	(0.00028)	(0.0004)	(0.004)	(0.0002)
I(DTB Law)	-945.23***	-0.308***	-0.007***	0.001	0.006***	-0.002	-0.001
	(319.22)	(0.088)	(0.002)	(0.001)	(0.002)	(0.016)	(0.001)
Relative Years to DTB Law*I(DTB Law)	-133.55**	-0.050**	-0.0008	-0.00003	0.0009**	0.0001	-0.0004**
	(66.11)	(0.023)	(0.0005)	(0.00027)	(0.0004)	(0.0035)	(0.0002)
% Effect At 10 Years	-4.20%	-2.07%	-1.83%	1.23%	12.30%	-0.01%	-0.81%
Panel B: Women							
	Earnings (i)	Hours Worked (ii)	Employed (iv)	Un-Employed (v)	Not in Labor Force (v)	Years of Education (vi)	Occup. Skill (vii)
Relative Years to DTB Law	-70.08	-0.062	-0.001	0.0000	0.001	0.001	-0.0005**
	(43.17)	(0.040)	(0.001)	(0.0003)	(0.001)	(0.004)	(0.0002)
I(DTB Law)	133.41	-0.057	-0.004	0.004**	0.0005	0.006	0.002
	(263.52)	(0.163)	(0.003)	(0.001)	(0.0037)	(0.022)	(0.001)
Relative Years to DTB Law*I(DTB Law)	57.43	-0.001	-0.001	-0.00001	0.001	0.001	0.0006***
	(43.80)	(0.040)	(0.001)	(0.00031)	(0.001)	(0.004)	(0.0002)
% Effect At 10 Years	2.33%	-0.23%	-1.92%	8.13%	4.73%	0.12%	1.43%

Notes: Authors' estimation of equation (2) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Relative Years to DTB Law is the number of years relative to the passage of a duty-to-bargain law when each cohort was 6 years old, which is set to zero for states that never pass such a law. I(DTB Law) is an indicator for whether a duty-to-bargain law has been passed in the state when each cohort was 6 years old. Regressions are based on 6,000 birth state-birth cohort-year observations. All estimates include birth state, year and birth cohort-by-year fixed effects as well as controls for racial/ethnic composition of the state-cohort-year-gender cell, exposure to school finance reform, food stamps and EITC when of school age. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. % Effect At 10 Years shows the calculated effect 10 years post DTB passage divided by the mean of each dependent variable. Standard errors clustered at the birth state level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 3.7: P-Values of Permutation Tests At 10 Years for Men

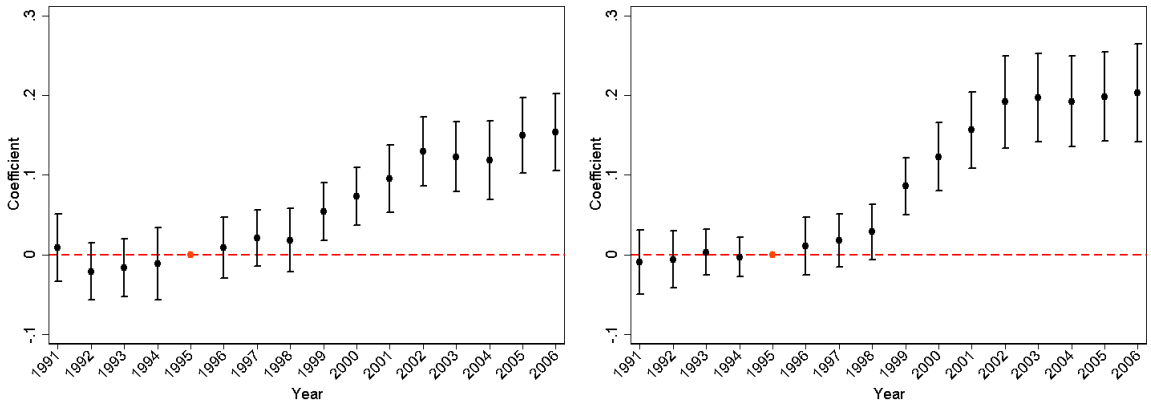
Panel A: Randomly Assigning Passage Dates						
	Earnings (i)	Hours Worked (ii)	Employed (iv)	Not in Labor Force (v)	Years of Education (v)	Occup. Skill (vi)
% Less than Baseline	0.003	0.073	0.063	0.947	0.113	0.080
Panel B: Randomly Assigning Passage Dates to Match Passage Timing Distribution						
	Earnings (i)	Hours Worked (ii)	Employed (iv)	Not in Labor Force (v)	Years of Education (v)	Occup. Skill (vi)
% Less than Baseline	0.000	0.027	0.033	0.993	0.080	0.093

Notes: All estimates include birth state, year and birth cohort-by-year fixed effects, as well as controls for racial/ethnic composition of the state-cohort-year-gender cell, exposure to school finance reforms, average state EITC and average food stamp availability during school years. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. The table shows the proportion of times the estimates from the permutation tests on the 10-year estimate are smaller than the baseline estimate. In Panel (A), we run 300 simulations in which we randomly assign passage dates to states. In Panel (B), we randomly assign passage dates to states in a way that matches the overall date-of-passage distribution shown in Figure 1.

Table 3.8: The Effect of Teacher Collective Bargaining on Male Non-Cognitive Skill Measures, NLSY79

	Rotter Locus of Control	Rosenberg Self-Esteem Scale	Pearlin Mastery Scale
I(DTB Law)	0.147	0.316	-0.148
	(0.203)	(0.489)	(0.249)
Years Post DTB Law	0.094**	-0.135**	-0.020
	(0.034)	(0.063)	(0.026)
Mean	8.41	22.68	22.29
% Effect at 10 Years	12.9%	-4.6%	-1.6%

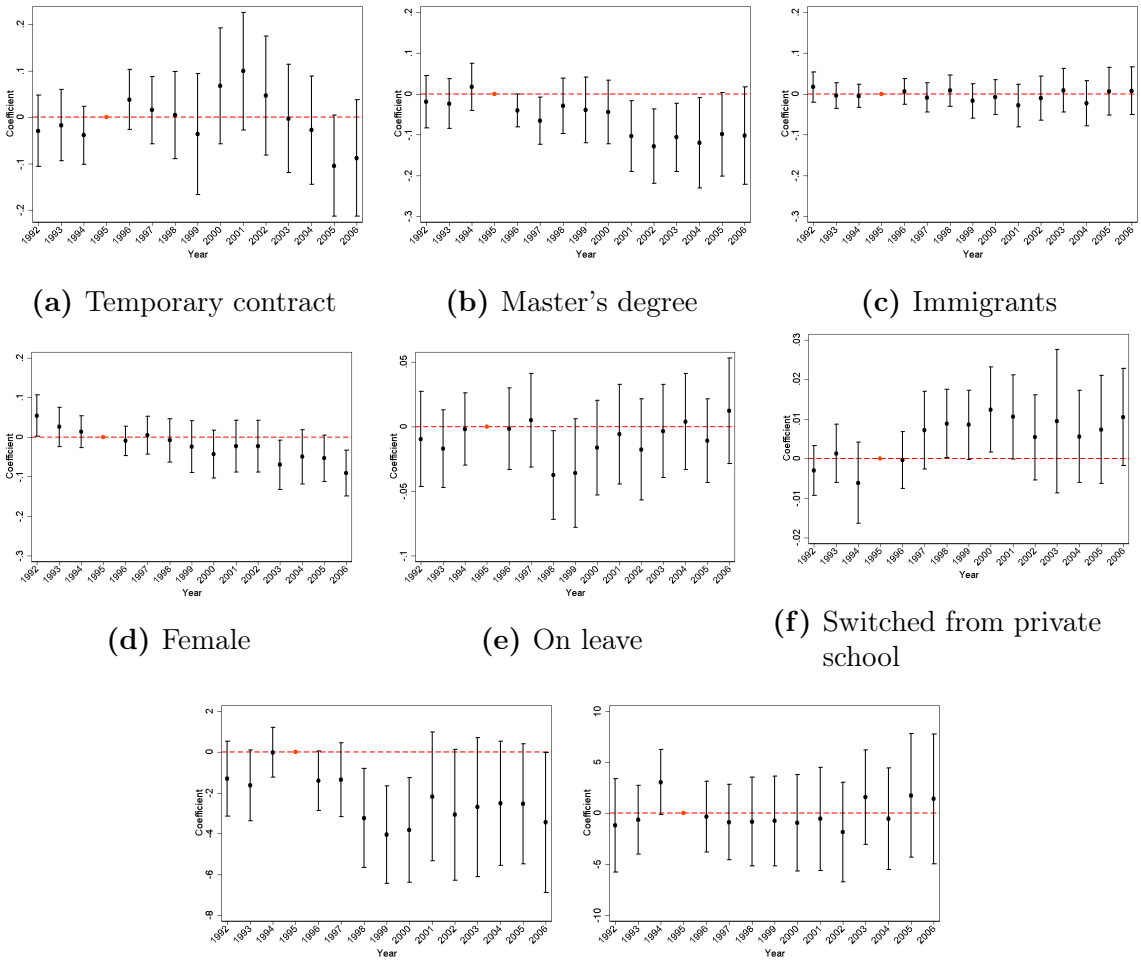
Notes: Data come from men in the NLSY79, 1957-1965 birth cohorts. All outcomes are measured in 1979. Models include controls for race and family income, mother's and father's education, as well as state at age 14 and birth cohort fixed effects. All estimates are weighted by the NLSY79 sample weights. The Rotter Locus of Control measures the extent to which students believe they have control over their lives: higher scores indicate less internal control (i.e., self-determination). The Rosenberg Self-Esteem Scale measures questions of self-worth, with higher scores associated with higher self-esteem. The Pearlin Mastery Scale measures the extent to which individuals perceive themselves in control of forces that significantly impact their lives, with higher scores indicating more control. Standard errors are clustered at the state level: ** indicates significance at the 5% level and * indicates significance at the 10% level.



(a) Male **(b)** Female

Figure A.1: Event study estimates by gender - Mean wage

Notes: Author's estimation of equation (2) as described in the text using 1991-2006 teacher registry data on the indicated subsample of public elementary school teachers in Sweden. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include year and municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.



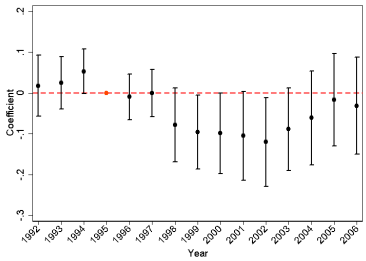
(a) Temporary contract **(b)** Master's degree **(c)** Immigrants

(d) Female **(e)** On leave **(f)** Switched from private school

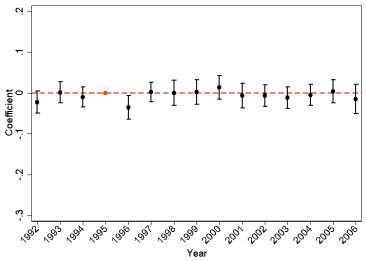
(g) Age **(h)** Percent worked

Figure A.2: Event study estimates - Labor supply outcomes

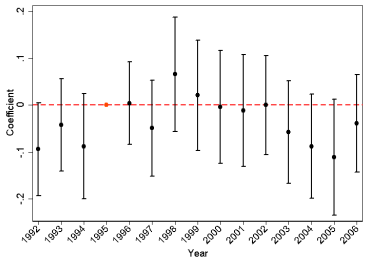
Notes: Author's estimation of equation (2) as described in the text using 1991-2006 teacher registry data on all public elementary school teachers in Sweden supplemented with 1992-2006 public-use data from Statistics Sweden and SNAE. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.



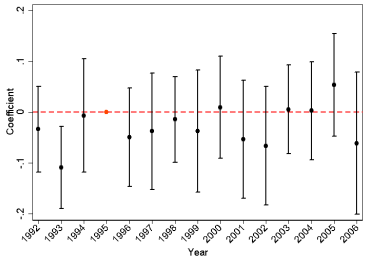
(a) Certificate



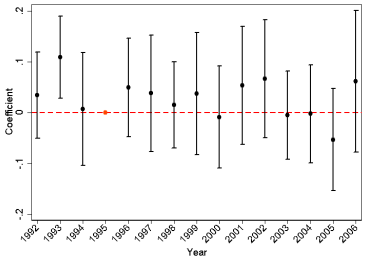
(b) Movers



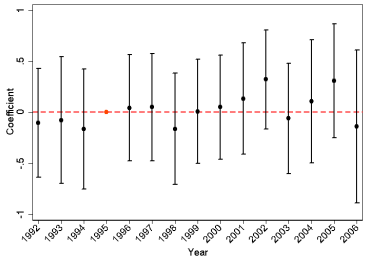
(c) Hires



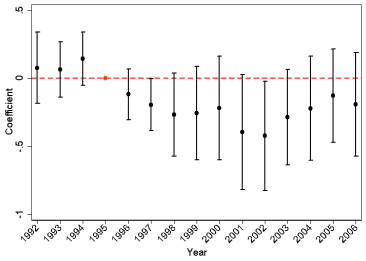
(d) Stayers



(e) Leavers



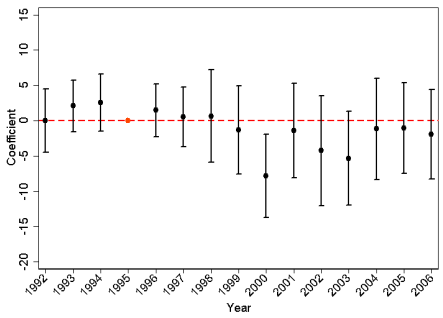
(f) High Movers



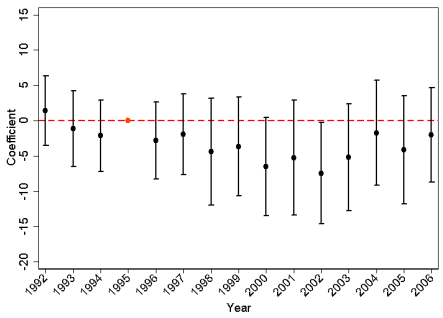
(g) Years of schooling

Figure A.3: Event study estimates - Labor supply outcomes

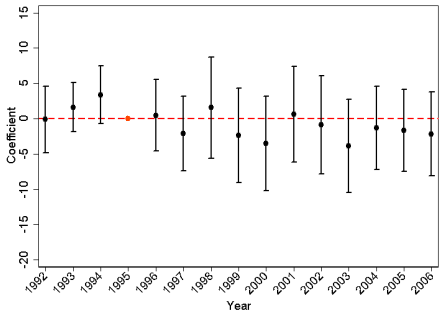
Notes: Author's estimation of equation (2) as described in the text using 1991-2006 teacher registry data on all public elementary school teachers in Sweden supplemented with 1992-2006 public-use data from Statistics Sweden and SNAE. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.



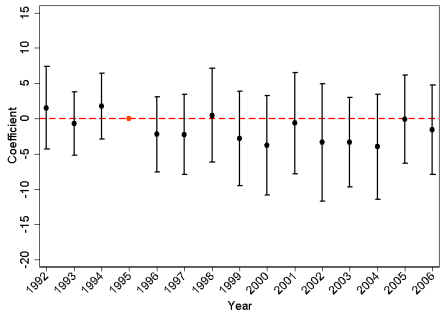
(a) GPA



(b) Math



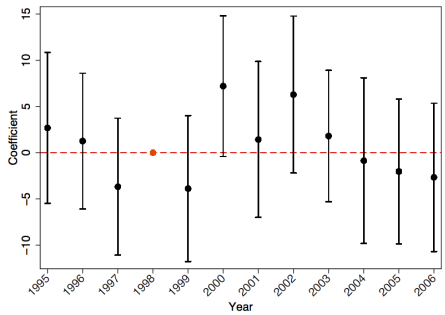
(c) Swedish



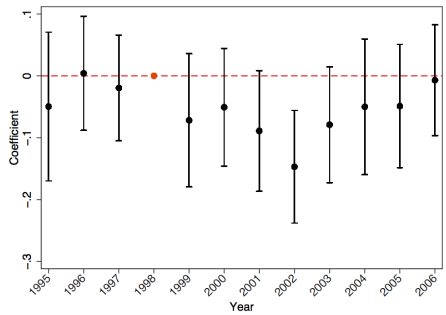
(d) English

Figure A.4: Event study estimates - Year 9 student outcomes

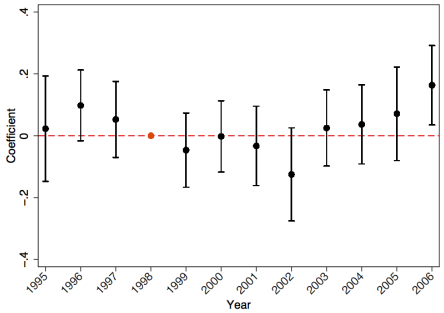
Notes: Author's estimation of equation (2) as described in the text using 1991-2006 registry data supplemented with public-use data from Statistics Sweden and SNAE. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.



(a) GPA



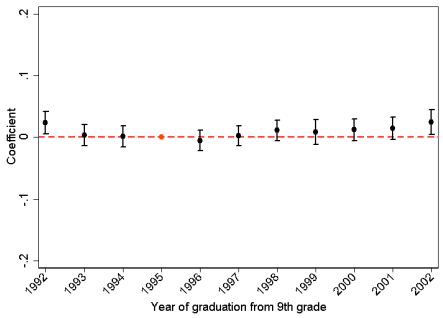
(b) Natural Science Track



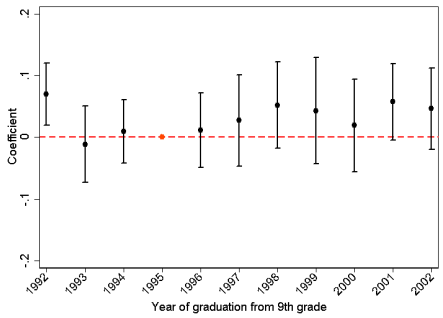
(c) University-Prep. Program

Figure A.5: Event study estimates - High school student outcomes

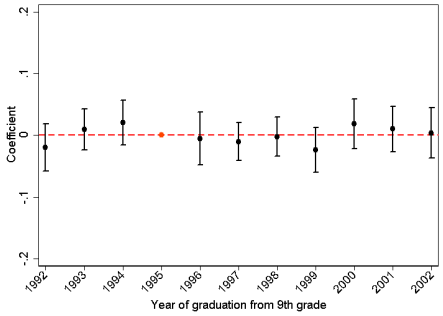
Notes: Author's estimation of equation (2) as described in the text using 1991-2006 registry data supplemented with public-use data from Statistics Sweden and SNAE. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.



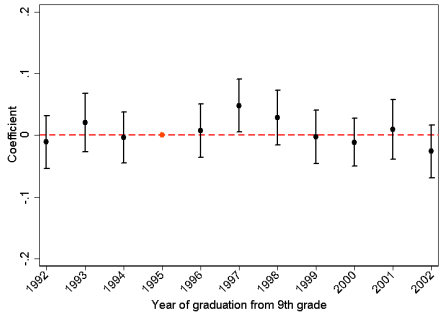
(a) Employment Income



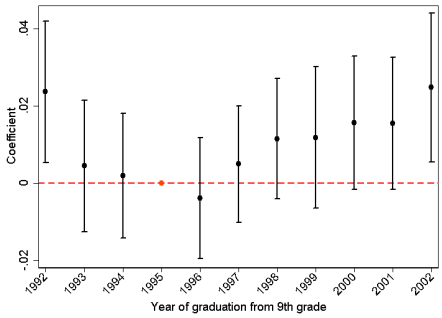
(b) Employment Sample



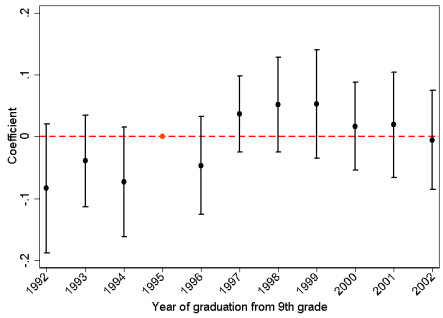
(c) Social Security



(d) Social Insurance Benefits



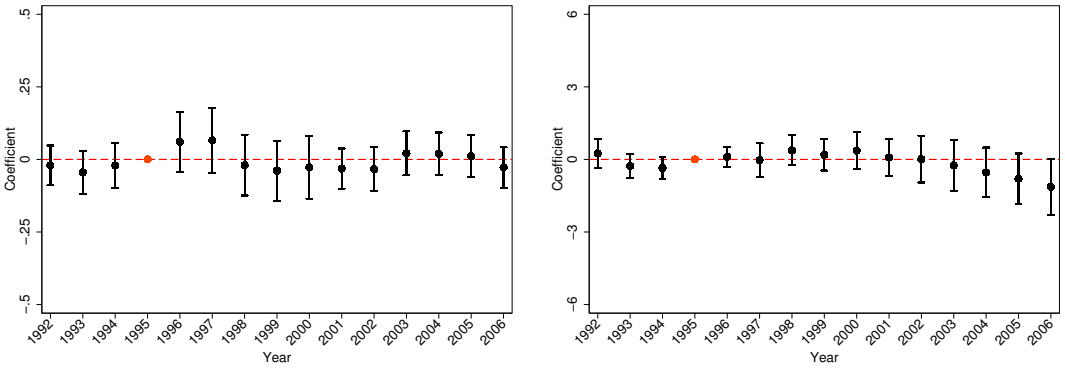
(e) Government-Funded Benefits



(f) University Enrollment

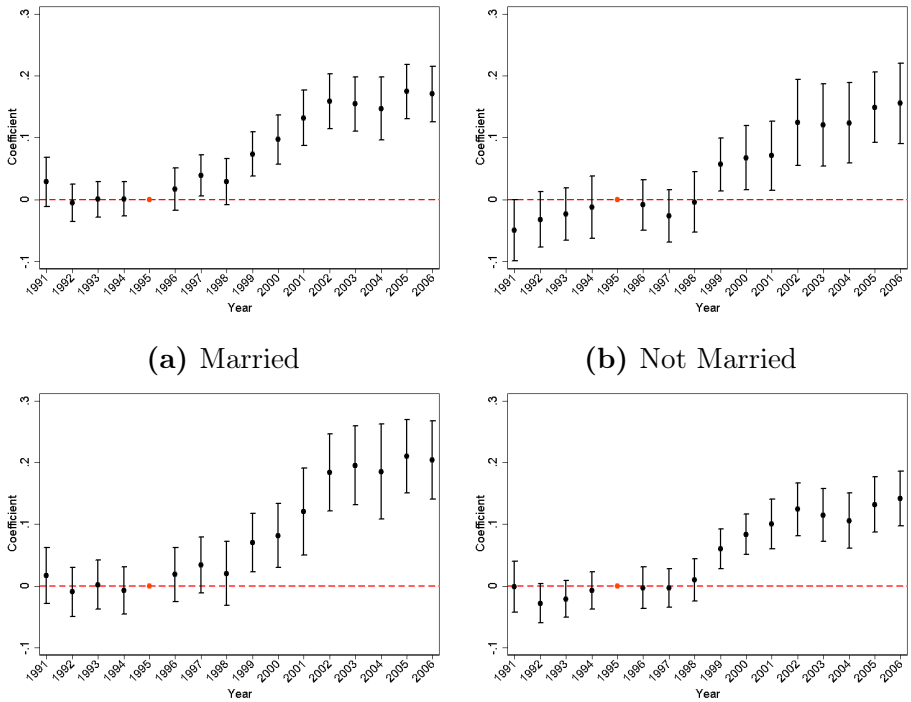
Figure A.6: Event study estimates - Higher education and labor market outcomes

Notes: Author's estimation of equation (2) as described in the text using 1991-2012 registry data supplemented with public-use data from Statistics Sweden and SNAE. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.



(a) Local tax rate **(b)** Teachers per 100 students

Figure A.7: Event study estimates - Local tax rate and teacher-student ratio
Notes: Author's estimation of equation (2) as described in the text using 1991-2006 registry data supplemented with public-use data from Statistics Sweden and SNAE. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.

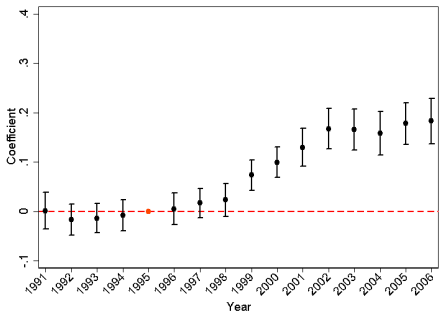


(a) Married **(b)** Not Married

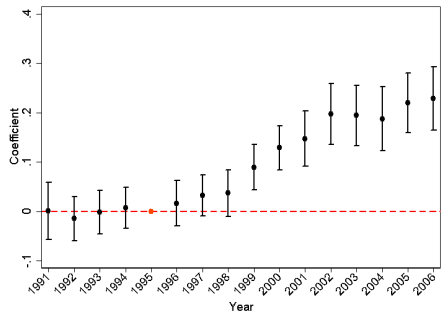
(c) Child present **(d)** No child present

Figure A.8: Event study estimates by teacher subgroups - Mean wage

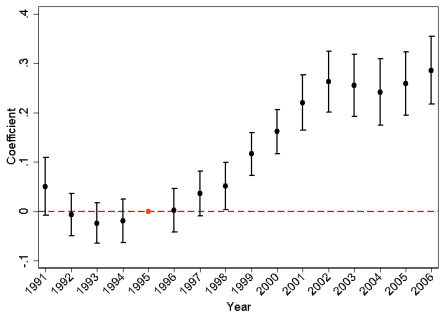
Notes: Author's estimation of equation (2) as described in the text using 1991-2006 teacher registry data on the indicated subsample of public elementary school teachers in Sweden supplemented with 1992-2006 public-use data from Statistics Sweden and SNAE. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.



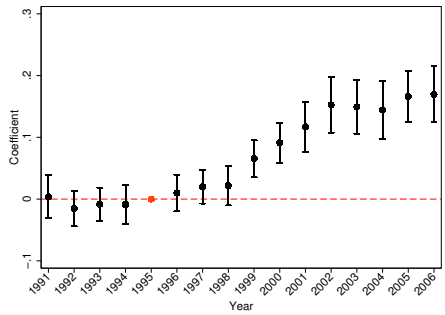
(a) Treatment based on five year pre-reform average non-teacher employment income



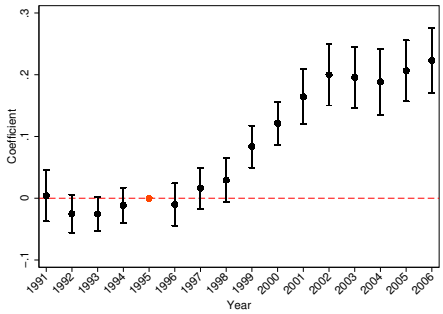
(b) Treatment based on middle 90 percent of non-teacher wage distribution



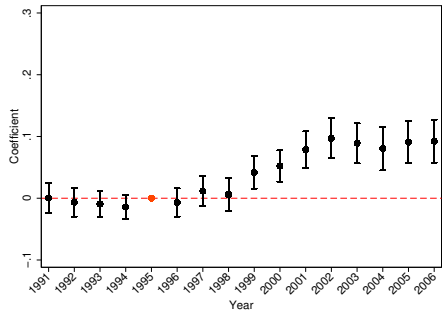
(c) Treatment based on predicted pre-refom teacher employment income



(d) Excluding municipalities that experienced border changes



(e) Treatment based on workers without college degrees



(f) Treatment defined on the municipality level

Figure A.9: Event study estimates - Sensitivity and robustness analyses

Notes: Author's estimation of equation (2) as described in the text using 1991-2006 teacher registry data on all public elementary school teachers in Sweden. Relative year -1 is omitted, so all estimates are in relationship to this year. All estimates include gender-by-year and gender-by-municipality fixed effects as well as municipality controls for fraction males, fraction immigrants, average income, fraction unemployed, average years of schooling, fraction on social security benefits and average age. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the municipality level.

Table A.1: Dependent variable sample means

	Mean
Panel A: Public elementary school teachers	
Switch from private school	0.002
On leave	0.022
Stayer	0.843
Leaver	0.157
Hire	0.145
Master	0.255
Age	46.120
Certificate	0.859
Temporary contract	0.161
Percent worked	91.994
Immigrant	0.075
Female	0.726
Mover	0.011
High mover	0.468
Log mean wage	9.946
Mean monthly wage (000 dollars)	2.640
Standard deviation (000 dollars)	0.321
Interquartile range (000 dollars)	0.477
Panel B: Student education and labor market outcomes	
Percentile math ranking, 9th grade	49.117
Percentile swedish ranking, 9th grade	48.903
Percentile english ranking, 9th grade	48.519
Percentile GPA ranking, 9th grade	48.985
University-preparatory program, high school	0.495
Science track, high school	0.180
GPA, high school	49.648
Ever enrolled, university	0.268
Log employment income, labor market	11.735
Employment sample, labor market	0.896
Social security recepient, labor market	0.046
Log social insurance benefits, labor market	5.915
Log government-funded benefits, labor market	9.939
Panel C: College-educated non-teachers	
Mean monthly employment income (000 dollars)	2.806

Notes: Author's calculation using 1992-2006 teacher registry data on all public elementary school teachers in Sweden, 1992-2006 educational attainment data from the grade 9 registry, 1995-2009 educational attainment data from the high school registry and 2002-2012 labor market data from LOUISE. Each observation is a municipality-gender-year. The number of observations are 8550 for all summary statistics except for the labor market and university outcomes in Panel B, which are based on 6270 observations. The reason is that these outcomes are measured ten years after 9th grade graduation, and 2012 is the most recent LOUISE data year I have access to. Thus, cohorts that graduated between 2003 and 2006 are excluded from these calculations.

Table A.2: Education Spending by Input

Input	Per Student Spending (000 dollars)	Fraction of Total Per Student Spending
Teachers	3.847	0.510
Supplies	0.309	0.041
Health	0.171	0.023
Food	0.485	0.064
Facilities	1.578	0.210
Other	1.171	0.152

Notes: Author's own calculation based on 1992-2006 public data released by the Swedish National Agency of Education. Costs represents real 2005 values and are shown in 000 dollars.

Table A.3: The effect of wage decentralization on labor supply stratified by teacher age

Panel A: Teachers 20-34 years olds

	Age	Master's Degree	Immigrants	Female	On leave	Switched from private school	Temporary contract	Percent worked
lnNTW*Post	-0.612	0.055	-0.041	0.087	0.031*	0.011	-0.092	4.244
	(0.854)	(0.078)	(0.041)	(0.061)	(0.017)	(0.010)	(0.098)	(4.947)
	Certificate	Certificate, alternative spec.	Movers	High Movers	Stayers	Leavers	Hires	Years of Schooling
lnNTW*Post	-0.218**	-0.194**	0.003	-0.034	0.007	-0.007	0.170**	-0.248
	(0.010)	(0.095)	(0.043)	(0.184)	(0.080)	(0.080)	(0.078)	(0.337)

Panel B: Teachers 35-49 years old

	Age	Master's Degree	Immigrants	Female	On leave	Switched from private school	Temporary contract	Percent worked
lnNTW*Post	-0.221	-0.051	0.005	-0.135***	-0.013	0.016***	0.053	-0.499
	(0.577)	(0.050)	(0.032)	(0.040)	(0.012)	(0.004)	(0.048)	(2.065)
	Certificate	Certificate, alternative spec.	Movers	High Movers	Stayers	Leavers	Hires	Years of Schooling
lnNTW*Post	-0.128***	-0.122**	0.002	0.324	0.034	-0.034	0.047	-0.359**
	(0.045)	(0.057)	(0.007)	(0.306)	(0.043)	(0.043)	(0.035)	(0.164)

Panel C: Teachers 50-64 years old

	Age	Master's Degree	Immigrants	Female	On leave	Switched from private school	Temporary contract	Percent worked
lnNTW*Post	1.383**	-0.155**	0.024	-0.012	0.015	0.004*	-0.021	-1.088
	(0.693)	(0.075)	(0.032)	(0.051)	(0.020)	(0.002)	(0.038)	(2.366)
	Certificate	Certificate, alternative spec.	Movers	High Movers	Stayers	Leavers	Hires	Years of Schooling
lnNTW*Post	-0.002	0.014	-0.001	0.000	-0.036	0.036	-0.027	-0.415**
	(0.036)	(0.047)	(0.006)	(0.400)	(0.036)	(0.036)	(0.025)	(0.196)

Notes: Author's estimation of equations (1) as described in the text using 1992-2006 teacher registry data on all public elementary school teachers in Sweden supplemented with 1992-2006 public-use data from Statistics Sweden and SNAE. Regressions are based on 8550 municipality-gender observations. All estimates include municipality-gender and year-gender fixed effects, as well as controls for gender composition, average years of schooling, fraction of social security benefit recepients, mean income, average age, fraction immigrants, fraction of student that attend private schools and total number of elementary school students. Standard errors clustered at the municipality level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table A.4: The effect of wage decentralization on teacher-student ratio and the local tax rate

	Teacher-student ratio	Local income tax rate
lnNTW*Post	0.010	0.003
	(0.361)	(0.034)

Notes: Author's estimation of equations (1) as described in the text using 1992-2006 registry data supplemented with public-use data from Statistics Sweden and SNAE. All estimates include municipality-gender and year-gender fixed effects, as well as controls for gender composition, average years of schooling, fraction of social security benefit recepients, mean income, average age and fraction immigrants. Standard errors clustered at the municipality level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table A.5: Public sector occupation groups that teachers came from, and left for, in the year prior to the reform

Occupation	Fraction of Teachers
Travel and Tourist Services	0.001
Research and Development within the Social Sciences and Humanities	0.001
Cleaning and Chimney Sweeping	0.001
Electric Power Supply Services	0.002
Transportation Support Services	0.002
Railroad Transportation	0.003
Cultural and Entertainment Services	0.003
Renting out Properties	0.004
Construction and Building Operations	0.006
Research and Development within the Natural Sciences and Technology	0.007
Employment and Recruitment Services	0.008
Foreign Affaris, Defense, Law Enforcement and Fire Protection	0.012
Religious and Interest Group Services	0.017
Sport Services	0.020
Library, Archive and Museum Services e.tc.	0.021
Other Recreational Services	0.044
Health Care Services	0.083
Public Administration	0.134
Social Services	0.631

Notes: Author's own calculation based on information from the Teacher Registry and the Wage Registry for Public Sector Employees between 1994 and 1995. Occupation groups are based on the Swedish three-digit SNI 92 classification.

Table A.6: Cross-LLM variation in annual pre-reform college-educated non-teacher employment income (000 dollars)

	Mean	Standard deviation
Panel A: Males		
College-educated non-teacher employment income, 1995	39.954	3.746
Panel B: Females		
College-educated non-teacher employment income, 1995	25.205	1.728

Notes: Author's calculation using information on employment income for all employed college-educated individuals in Sweden from the 1995 Longitudinal Database for Education, Income and Labor Market Participation. Values have been converted to represent real 2005 dollars. See data section for details on sample construction.

Table A.7: The effect of wage decentralization on wage structure

	Mean Wage	10th Percentile Wage	90th Percentile Wage
lnNTW*Post (Decile 1)	0.0092	0.0022	0.010
	(0.0066)	(0.0126)	(0.0084)
lnNTW*Post (Decile 2)	0.0332***	0.0269	0.0301*
	(0.0113)	(0.0226)	(0.0162)
lnNTW*Post (Decile 3)	0.0780***	0.0833**	0.0611***
	(0.0160)	(0.0334)	(0.0205)
lnNTW*Post (Decile 4)	0.0922***	0.109***	0.0655***
	(0.0164)	(0.0319)	(0.0192)
lnNTW*Post (Decile 5)	0.1020***	0.1370***	0.0634***
	(0.0159)	(0.0318)	(0.0194)
lnNTW*Post (Decile 6)	0.1050***	0.1410***	0.0621***
	(0.0148)	(0.0297)	(0.0178)
lnNTW*Post (Decile 7)	0.106***	0.133***	0.0609***
	(0.0146)	(0.0281)	(0.0181)
lnNTW*Post (Decile 8)	0.100***	0.115***	0.0704***
	(0.0136)	(0.0264)	(0.0148)
lnNlnNTW*PostTW (Decile 9)	0.0765***	0.1000***	0.0459***
	(0.0133)	(0.0242)	(0.0140)

Notes: Author's estimation of equations (1) as described in the text using 1991-2006 teacher registry data on all public elementary school teachers in Sweden. Regressions are based on 9120 municipality-gender observations. All estimates include municipality-gender and year-gender fixed effects, as well as controls for gender composition, average years of schooling, fraction of social security benefit recepients, mean income, average age and fraction immigrants. Standard errors clustered at the municipality level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

The Structural Break Method

This method is similar to that of identifying breaks in time series data, and consists of estimating the following regression

$$Dn_{s,m,t} = \alpha_m + d_m \mathbf{1}[i_{s,m,t-10} > i^*_{m,t-10}] + \varepsilon_{s,m,t}, \qquad \text{for } 0 \leq i_{s,m,t-10} \leq I$$

where $Dn_{s,m,t} = \frac{N_{s,m,t} - N_{s,m,t-10}}{P_{s,m,t-10}}$ and represents the change in the native population in neighborhood *s* in metropolitan area *m* between *t-10* and *t*, and $d_m \mathbf{1}[i_{s,m,t-10} > i^*_{m,t-10}]$ is an indicator variable that takes the value of one if the immigrant share in the neighborhood exceeds the tipping point of the metropolitan area.

To obtain estimates of the tipping points in the metropolitan areas, $i^*_{m,t-10}$, we restrict the tipping points to be in the interval *[0, 50%]* and choose the values that maximizes R^2 of the above equation, separately for each metropolitan area. According to Card et al. (2008), this method works well for identifying tipping points in large cities, but performs less well in small cities due to a tendency to identify tipping points that reflects clear outliers. Given the average size of the metropolitan areas in Sweden it is therefore inappropriate to rely on this strategy for the purpose of identifying the tipping points.

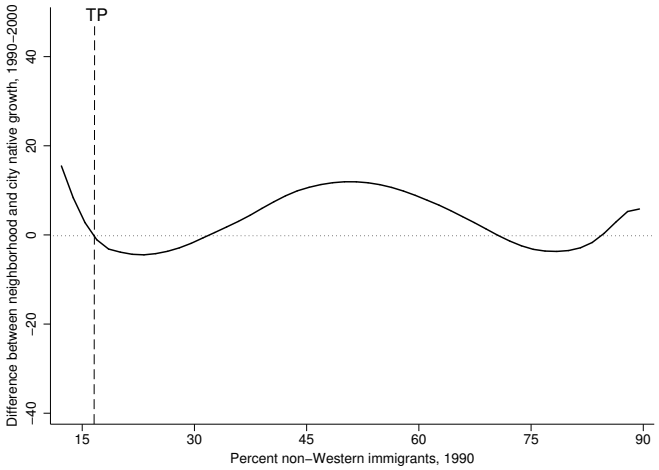
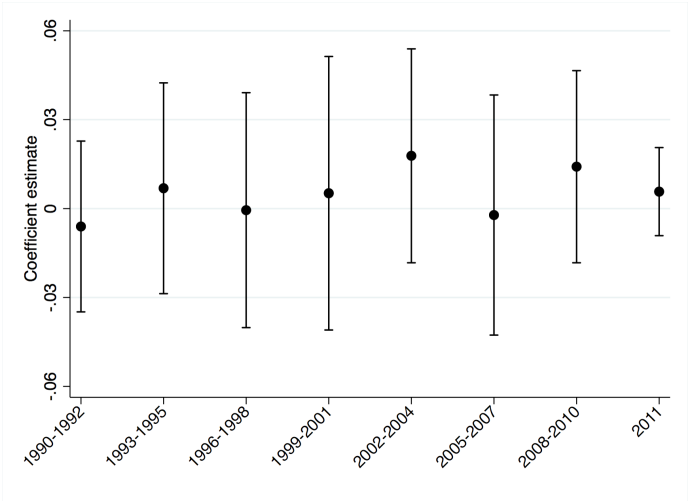
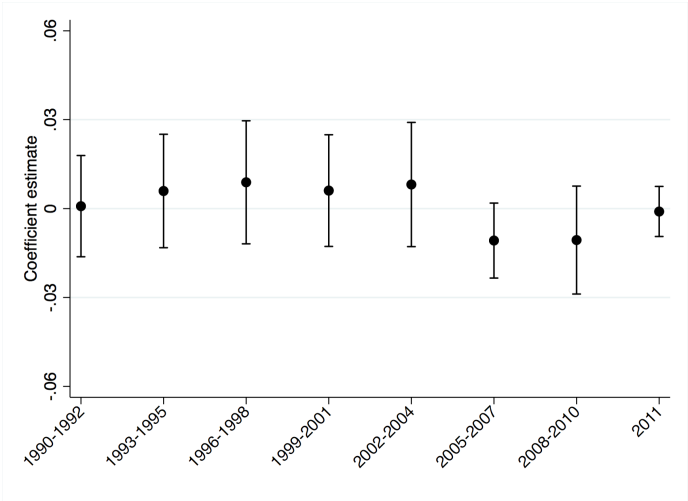


Figure B.1: Illustration of the search method for identifying the tipping point
graph1.eps

Notes: The figure demonstrates how the location of the tipping point is derived from equation (1) for a hypothetical city. The solid line depicts the growth function of neighborhood native population modelled as a fourth-order polynominal. The horizontal line shows where the dependent variable of equation (1) is equal to zero. The proposed tipping point is located at the intersection of this line and the growth function, denoted by the dashed vertical line. As illustrated in the Figure, and discussed in the text, there can be more than one root, and in such cases we follow Card et al. (2008) and pick the root associated with the most negative slope.



(a) Immigrants



(b) Natives

Figure B.2: Time heterogeneity in treatment effects

Notes: The unit of observation is an individual born between 1948 and 1958 (Old Cohort) that resided in one of the 520 neighborhoods included in our analysis in the base year not used to identify the location of the tipping points. The figure depicts the point estimates obtained from estimating equation (4) seperateley on three year averages of employment income, stratified by nativity status. All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. Demographic controls are gender, educational attainment, income and binaries for whether this information was not available for the individual, all measured in the base year. All models include birth year and municipality fixed effects. Natives refer to individuals not born in, and do not have a parent born in, a non-Western country. The bars depict the 95% confidence intervals associated with each point estimate.

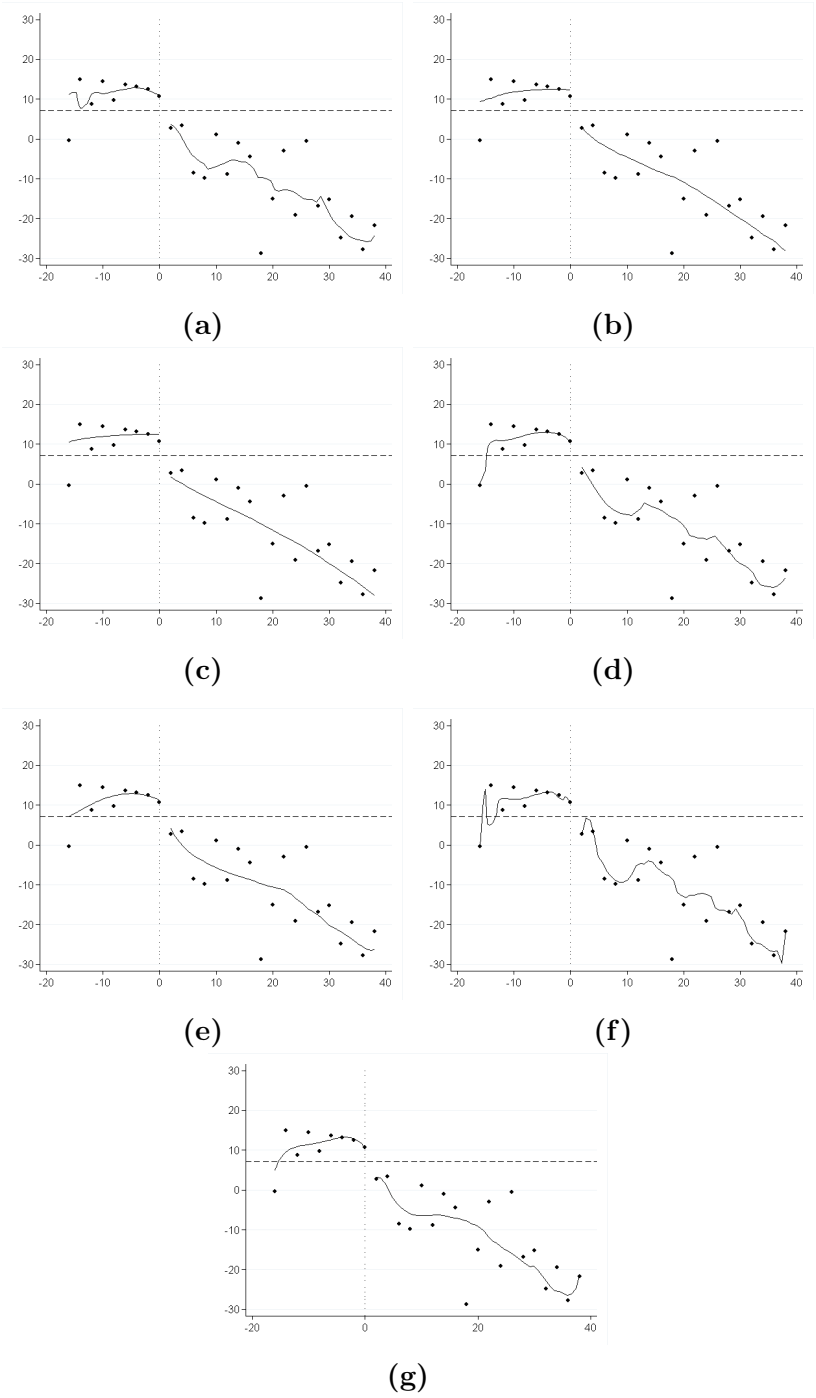


Figure B.3: Discontinuity in native population change around candidate tipping point, alternative bandwidths and degrees of smoothing

Notes: Dots show mean change in neighborhood native population growth between 1990 and 2000, grouping neighborhoods into 2% bins by the deviation in immigrant share from the estimated tipping point in the base year. The vertical lines depict the estimated tipping points (normalized to zero). The solid lines represent regressions fitted separately on either side of the tipping point weighted by the size of the neighborhoods and the fraction of the decade that the neighborhood spent above the tipping point, using an Epanechnikov kernel. Only the 1/3 of the sample not used for identifying the location of the tipping points is used for these visual depictions.

Table B.1: Foreign-born by country of birth

Country	1950	1960	1970	1980	1990	2000	2010	2015
Panel A: Largest source countries 2015								
Finland	44,821	101,307	235,453	251,342	217,636	195,447	169,521	156,045
Iraq	5	16	108	631	9,818	49,372	121,761	131,888
Syria	0	6	100	1,606	5,874	4,162	20,758	98,216
Poland	7,832	6,347	10,851	19,967	35,631	40,123	70,253	85,517
Iran	110	115	411	3,348	40,048	51,101	62,120	69,067
Yugoslavia	171	1,532	33,779	37,982	43,346	71,972	70,819	67,190
Somalia	0	0	16	100	1,441	13,082	37,846	60,623
Bosnia and Herzegovina	0	0	0	0	0	51,526	56,183	57,705
Germany	21,652	37,580	41,793	38,974	37,558	38,155	48,158	49,586
Turkey	87	202	3,768	14,357	25,528	31,894	42,527	46,373
Panel B: Largest source countries 1950								
Finland	44,821	101,307	235,453	251,342	217,636	195,447	169,521	156,045
Norway	31,312	37,253	44,681	42,863	52,744	42,464	43,430	42,047
Estonia	25,062	*	18,513	15,331	11,971	10,253	10,010	10,303
Denmark	22,801	35,112	39,152	43,501	43,931	38,190	45,584	41,870
Germany	21,652	37,580	41,793	38,974	37,558	38,155	48,158	49,586
United States	10,713	10,874	12,646	11,980	13,001	14,413	17,179	19,515
Poland	7,832	6,347	10,851	19,967	35,631	40,123	70,253	85,517
Latvia	4,423	*	3,244	2,664	2,025	2,305	4,686	7,026
Czechoslovakia	3,548	3,562	7,392	7,529	8,432	7,304	5,970	5,293
Austria	2,665	5,809	7,927	6,995	6,530	6,021	5,829	5,772
Panel C: Source countries by continents								
The nordic countries	99,080	174,043	320,913	341,253	319,082	279,631	263,227	245,633
EU25 (excl. the nordic countries)	75,631	75,138"	137,251	148,459	164,961	172,599	274,247"'	331,026
Rest of Europe	1,766	4,048	43,104	57,292	81,885	189,766	215,975"'	238,565
Africa	355	596	4,149	10,025	27,343	55,138	114,853	178,624
North America	11,334	11,665	15,629	14,484	19,087	24,312	31,263	35,780
South America	412	679	2,300	17,206	44,230	50,853	63,725	68,571
Asia	905	1,476	5,949	30,351	124,447	220,677	410,083	565,050
Oceania	93	211	558	962	1,866	2,981	4,529	5,245
Unknown	137	162	488	97	73	257	818	1,148
Panel D: Non-Western foreign-born								
Non-Western	48,904	30,070	130,804	201,373	380,945	623,042	991,482	1,285,961
Panel E: Total immigration								
Total Foreign-born	197,810	229,879	537,585	626,953	790,445	1,003,798	1,384,929	1,676,264
Percent Foreign-born	2.8	3.1	6.7	7.5	9.2	11.3	14.7	17.0
Total Population	7,041,829	7,495,129	8,076,903	8,317,235	8,590,630	8,882,792	9,415,570	9,851,017

Notes: * Included in the calculation of Soviet Union immigrants; ' Including Estonia, Latvia and Lithuania; " Excluding Estonia, Latvia and Lithuania; "'Calculation based on EU28. Source: Authors' calculations based on data from Statistics Sweden (2015).

Table B.2: Neighborhood crossovers

Year of tipping	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Neighborhoods	156	16	9	8	9	5	6	4	4	1

Notes: The tables shows the number of neighborhoods in which the share of non-Western immigrants increased from being below the candidate tipping point in 1990 to being above the treshold for each year between 1991 and 1999. The table further shows the number of neighborhoods that had immigrant shares above the identified tipping pooint in 1990. The sample used is the 1/3 sample not used for identifying the location of the tipping points.

Table B.3: Donut-style regression discontinuity models for changes in native population around candidate tipping points

	Change in native population				
	0.10 Donut Hole	0.3 Donut Hole	0.5 Donut Hole	1.00 Donut Hole	2.00 Donut Hole
	(i)	(ii)	(iii)	(iv)	(v)
Beyond TP	-0.095**	-0.093**	-0.096**	-0.104**	-0.109**
	(0.039)	(0.042)	(0.042)	(0.046)	(0.049)
Observations	517	514	511	501	488

Notes: The unit of observation is a neighborhood as identified by the SAMS code. Results are obtained from estimating equation (2). Across the columns, neighborhoods with base year immigrant shares +/- 0.05 (i), 0.15 (ii), 0.25 (iii), 0.50 (iv) and 1.00 (v) of the identified tipping point are excluded from the estimation. Years of treatment has been instrumented by whether the neighborhood was above or below the tipping point in the base year. All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. The sample used for estimation is the 1/3 sample not used for identifying the tipping points. Demographic controls are years of schooling, income and gender, all measured in the base year. The regressions are weighted by the size of the neighborhoods. All specifications include metropolitan area fixed effects. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table B.4: Regression discontinuity models for population changes around candidate tipping points, Western immigrants

	Native Growth	Western Immigrant Growth	Non-Western Immigrant Growth
Beyond TP	-0.027	-0.002	-0.003
	(0.060)	(0.007)	(0.027)
Observations	520	520	520

Notes: The unit of observation is a neighborhood as identified by the SAMS code. The results are obtained from estimating new tipping points based on fraction Western immigrants using equation (1), and then using these new candidate thresholds to estimate equation (2). All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. The sample used for estimation is the 1/3 sample not used for identifying the location of the tipping points. Demographic controls are years of schooling, income and gender, all measured in the base year. The regressions are weighted by the size of the neighborhoods. All specifications include metropolitan area fixed effects. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table B.5: Sensitivity analysis on the change in native population growth around the candidate tipping point

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Beyond TP	-0.088**	-0.092**	-0.109***	-0.091**	-0.091**	-0.089*	-0.091**	-0.091**	-0.082***
	(0.029)	(0.040)	(0.039)	(0.041)	(0.039)	(0.048)	(0.042)	(0.039)	(0.036)
Polynomial	Linear	Quadratic	Cubic	Quartic	Quartic	Quintic	Quartic	Quartic	Quartic
Baseline Controls	x	x	x		x	x	x	x	x
Fully Interacted							x		
Additional Controls								x	
Population Density									x
Observations	520	520	520	520	520	520	520	520	520
R-squared	0.287	0.287	0.302	0.233	0.305	0.305	0.305	0.334	0.313

Notes: The unit of observation is a neighborhood as identified by the SAMS code. The results are obtained from estimating equation (2). Dependent variable is change in native population between 1990 and 2000. Standard errors are clustered on one percent bins of the running variable. The sample used for estimation is the 1/3 sample not used for identifying the location of the tipping points. Baseline controls are years of schooling, income and gender, all measured in the base year. Additional controls are years since migration, number of children in household and social welfare recipient status. The regressions are weighted by the size of the neighborhoods. All specifications include metropolitan area fixed effects. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table B.6: Regression discontinuity models for changes in residential population composition around candidate tipping points, local linear regression

	Native Growth	Immigrant Growth	Population Growth
Beyond TP	-0.112**	0.005	-0.107*
	(0.050)	(0.021)	(0.057)
R-squared	0.190	0.301	0.080
Observations	433	433	433

Notes: The unit of observation is a neighborhood as identified by the SAMS code. The bandwidth has been chosen using the cross-validation method proposed by Ludwig and Miller (2005). $h = 11.58483$. The sample used for estimation is the 1/3 sample not used for identifying the location of the tipping points. Demographic controls are years of schooling, income and gender, all measured in the base year. The regressions are weighted by the size of the neighborhoods. All specifications include metropolitan area fixed effects. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table B.7: Selective migration

		Educational Level		Income Level		Gender		Excluding
	Baseline	High	Low	High	Low	Men	Women	Outliers
Panel A: Young Cohort								
Control to Treatment	-0.053***	-0.034***	-0.071***	-0.040***	-0.070***	-0.049***	-0.052***	-0.048***
	(0.008)	(0.006)	(0.010)	(0.009)	(0.007)	(0.003)	(0.003)	(0.010)
Treatment to Control	0.283***	0.308***	0.264***	0.332***	0.258***	0.279***	0.248***	0.276***
	(0.020)	(0.030)	(0.019)	(0.021)	(0.022)	(0.023)	(0.011)	(0.021)
Observations	62,525	29,687	32,838	15,649	15,559	32,030	30,495	60,930
Panel B: Middle Cohort								
Control to Treatment	-0.010***	-0.073***	-0.120***	-0.088***	-0.114***	-0.096***	-0.115***	-0.085***
	(0.020)	(0.017)	(0.024)	(0.025)	(0.017)	(0.020)	(0.007)	(0.024)
Treatment to Control	0.264***	0.301***	0.240***	0.363***	0.218***	0.259***	0.258***	0.259***
	(0.014)	(0.016)	(0.015)	(0.026)	(0.016)	(0.015)	(0.009)	(0.014)
Observations	56,637	26,826	29,811	14,148	14,149	29,079	27,558	55,256
Panel C: Old Cohort								
Control to Treatment	-0.059***	-0.042***	-0.070***	-0.043***	-0.071***	-0.062***	-0.054***	-0.051***
	(0.009)	(0.006)	(0.012)	(0.012)	(0.010)	(0.009)	(0.003)	(0.010)
Treatment to Control	0.304***	0.393***	0.260***	0.392***	0.274***	0.313***	0.242***	0.299***
	(0.028)	(0.034)	(0.021)	(0.032)	(0.031)	(0.031)	(0.029)	(0.030)
Observations	92,798	36,437	56,361	23,195	23,183	47,314	45,484	91,221

Notes: The unit of observation is an individual that resided in one of the 520 neighborhoods not used to estimate the location of the tipping points. The results are obtained from estimating equation (3). All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. Demographic controls are gender, mother's education, father's education, parental income and binaries for whether this information was not available for the individual, all measured in the base year. For the Old Cohort, parental education and income have been replaced with own education and income. All models include birth year and metropolitan area fixed effects. Immigrants refer to individuals born in, or that have at least one parent born in, a non-Western country. Column (2) and (3) stratify the sample based on whether the individual has at least one parent with post-secondary education for the young and middle cohorts, and based on whether the individual has or does not have post-secondary education for the old cohort. Columns (4) and (5) stratify the sample based on whether the individual's parental income is in the bottom or top quartile of the income distribution for the youth and middle cohorts, and based on whether the individual is in the bottom or the top quartile of the income distribution for the old cohort. Columns (6) and (7) statify the sample based on gender. Column (8) exclude individuals from neighborhoods in the right-tail of the immigrant share distribution. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table B.8: The reduced-form effect of neighborhood composition on cognitive and non-cognitive military test scores

	Cognative	Non-cognative
Panel A: Immigrants		
i. 1973-1983		
Beyond TP	0.008	-0.070
	(0.198)	(0.196)
i. 1973-1980		
Beyond TP	-0.070	-0.169
	(0.242)	(0.212)
Panel B: Natives		
i. 1973-1983		
Beyond TP	-0.048	0.128
	(0.066)	(0.083)
i. 1973-1980		
Beyond TP	-0.017	0.183*
	(0.061)	(0.097)

Notes: The unit of observation is an individual that was started school 1973-1983 (Row 1) or 1973-1980 (Row 2) and resided in one of the 520 neighborhoods included in our analysis in the base year. The results are obtained from estimating equation (3). All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. Demographic controls are gender, mother's education, father's education, parental income and indicators for whether this information was not available for the individual. All models include birth year and metropolitan area fixed effects. Immigrants refer to individuals born in, or that have at least one parent born in, a non-Western country. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table B.9: Fraction of individuals that maintain treatment status over time

	All		Natives		Immigrants	
Year	Control	Treatment	Control	Treatment	Control	Treatment
1991	0.85	0.82	0.85	0.79	0.85	0.87
1992	0.79	0.73	0.79	0.70	0.77	0.78
1993	0.74	0.64	0.74	0.62	0.70	0.70
1994	0.68	0.58	0.68	0.55	0.64	0.64
1995	0.64	0.53	0.64	0.50	0.61	0.59
1996	0.59	0.49	0.59	0.46	0.56	0.54
1997	0.56	0.45	0.56	0.42	0.54	0.51
1998	0.52	0.41	0.52	0.39	0.50	0.47
1999	0.49	0.39	0.49	0.36	0.48	0.45
2000	0.48	0.37	0.47	0.35	0.47	0.43

Notes: The unit of observation is an individual that resided in one of the 520 neighborhoods included in our analysis in the base year. The Treatment columns depict the fraction of individuals that resided in a neighborhood with an immigrant share above the candidate threshold in the base year and remained in a neighborhood with an immigrant share above the threshold in year t. The Control columns depict the fraction of individuals that resided in a neighborhood with an immigrant share below the candidate threshold in the base year and remained in a neighborhood with an immigrant share below the threshold in year t.

Table B.10: Neighborhood population density

	All	Stockholm	Gothenburg	Malmo
Mean	4074.34	2437.36	5595.21	5326.58
S.D.	4535.89	2992.11	5390.15	4414.88

Notes: Authors' own calculations based on information on land size from Jan Amcoff and data from IFAU. See Amcoff (2012) for detailed information on how the density measure was constructed.

Table B.11: Tipping behavior of neighboring neighborhoods

	Mean	Standard Deviation	No Tipped Neighbors	All Neighbors Tipped	Number of Tipped Neighborhoods
All	0.62	0.35	0.12	0.25	459
Stockholm	0.43	0.29	0.17	0.08	166
Gothenburg	0.75	0.34	0.09	0.53	208
Malmo	0.65	0.31	0.09	0.24	85

Notes: Authors' own calculations using Statistic Sweden's SAMS Atlas. In a first step, neighborhoods with immigrant shares above the threshold in the base year are identified. In a second step, the SAMS Atlas is used to obtain the names of the neighborhoods surrounding the tipped neighborhoods. Finally,data from IFAU is used to identify the fraction of these neighborhoods that have tipped.

Table B.12: The reduced form effect of neighborhood composition on short-term labor market outcomes

	Self-Employment Income	Employment Income	Government-Funded Benefits
Panel A: Immigrants			
i. Intensive Margin			
Beyond TP	-0.220	0.026	0.015
	(0.171)	(0.045)	(0.068)
Observations	1803	16007	6704
ii. Extensive Margin			
Beyond TP	0.009	0.018	-0.007
	(0.006)	(0.020)	(0.014)
Observations	23253	23253	23253
Panel B: Natives			
i. Intensive Margin			
Beyond TP	-0.136	0.007	-0.020
	(0.127)	(0.023)	(0.054)
Observations	6315	81268	25528
i. Extensive Margin			
Beyond TP	0.006	-0.003	-0.017**
	(0.009)	(0.011)	(0.008)
Observations	93953	93953	93953

Notes: The unit of observation is an individual born between 1948 and 1958 that resided in one of the 520 neighborhoods included in our analysis that were not used to estimate the location of the tipping point . The results are obtained from estimating equation (3). All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. Demographic controls are gender, years of schooling, income and indicators for whether this information was not available for the individual. All models include birth year and metropolitan area fixed effects. All dependent variables are measured in 2000. All controls are measured in 1990. Immigrants refer to individuals born in, or that have at least one parent born in, a non-Western country. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table B.13: Descriptive statistics of neighborhoods included/excluded from analysis

	Included	Excluded
Fraction Natives	0.81(0.14)	0.82(0.21)
Fraction Females	0.49 (0.03)	0.43 (0.17)
Age	39.33 (2.97)	41.02 (6.19)
Years Since Migration	17.26 (3.92)	17.90 (8.12)
Fraction With University Education	0.10 (0.08)	0.08 (0.12)
Employment Income (000s SEK)	165.52 (46.81)	138.78 (70.34)
Fraction on Social Welfare	0.07 (0.09)	0.07 (0.16)
Native Growth Rate	0.09 (0.30)	2.48 (17.12)
Immigrant Growth Rate	0.07 (0.12)	0.91 (7.02)
Total Growth Rate	0.15 (0.33)	3.39 (23.80)

Notes: Authors' own calculations using population-wide registry data from IFAU. Values represent unweighted means, and standard deviations are provided in brackets. Salary refers to income from primary occupation, and includes zeros.

Table B.14: The effect of tipping on neighborhood environment

	Economic Activity Index	Sociodemographic Index
Neighborhood Analysis		
All	-0.073	-0.339***
	(0.056)	(0.101)
Individual-level Analysis		
All	-0.101	-0.251***
	(0.129)	(0.087)
Natives	-0.074	-0.288***
	(0.147)	(0.101)
Immigrants	-0.100	-0.201**
	(0.113)	(0.083)
Stayers	-0.228	-0.347***
	(0.164)	(0.126)
Leavers	0.068	-0.229**
	(0.111)	(0.087)

Notes: The unit of observation is an individual that resided in one of the 520 neighborhoods not used to estimate the location of the tipping points. The neighborhood analysis results are obtained from estimating equation (2), while the individual level analysis results are obtained from estimating equation (3). All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. Demographic controls are gender, mother's education, father's education, parental income and binaries for whether this information was not available for the individual, all measured in the base year. For the Old Cohort, parental education and income have been replaced with own education and income. All models include birth year and metropolitan area fixed effects. Immigrants refer to individuals born in, or that have at least one parent born in, a non-Western country. The Economic Activity Index is based on three labor market variables (average employment income, average education, and fraction employed) while the Sociodemographic Index is based on four sociodemographic variables (gender balance, age profile, fraction immigrants and fraction on social security benefits). For each of these indices, we use unity-based normalization to bring the values of each of the individual variables into the range [0,1], take their sum, and then standardize the index to have a mean of zero and a standard deviation of one.*** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table B.15: The reduced form effect of neighborhood composition on immigrants, sensitivity table

	9th grade GPA	9th grade Swedish	High School GDP	Years of schooling	Empl. Sample	Empl. Income
Panel A: Young Cohort						
Population Density	0.880	1.410	0.319	0.021	0.013	0.231
	(1.677)	(1.959)	(1.574)	(0.097)	(0.019)	(0.230)
Excl. Neighborhoods with 100% Tipped Neighbors	0.749	1.148	0.251	0.034	0.017	0.275
	(1.438)	(1.829)	(1.527)	(0.094)	(0.019)	(0.232)
Excluding Outliers	-2.194	1.111	0.056	-0.015	0.002	0.059
	(1.612)	(1.823)	(1.661)	(0.098)	(0.023)	(0.285)
Panel B: Middle Cohort						
Population Density	-0.353	-0.458	-0.396	0.153	0.003	0.080
	(1.220)	(2.039)	(1.950)	(0.149)	(0.020)	(0.251)
Excl. Neighborhoods with 100% Tipped Neighbors	-0.181	-0.530	-0.553	0.154	0.005	0.105
	(1.108)	(2.077)	(1.818)	(0.162)	(0.019)	(0.237)
Excluding Outliers	-1.247	0.041	-0.288	0.199	0.000	0.045
	(1.300)	(2.109)	(2.117)	(0.161)	(0.019)	(0.241)
Panel C: Old Cohort						
Population Density					0.011	0.111
					(0.026)	(0.331)
Excl. Neighborhoods with 100% Tipped Neighbors					0.020	0.232
					(0.028)	(0.353)
Excluding Outliers					0.009	0.125
					(0.029)	(0.376)

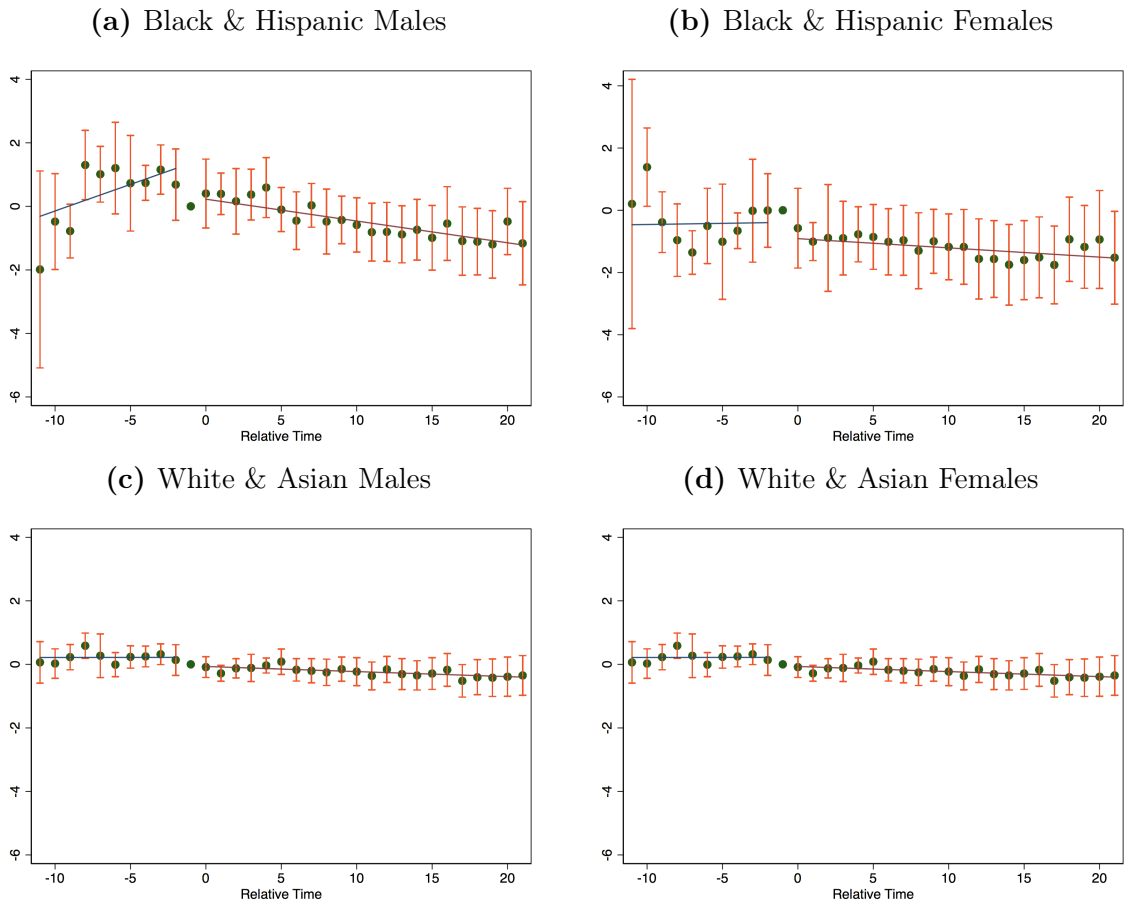
Notes: The unit of observation is an individual that resided in one of the 520 neighborhoods included in our analysis in the base year. The results are obtained from estimating equation (3). All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. Demographic controls are gender, mother's education, father's education, parental income and binaries for whether this information was not available for the individual, all measured in the base year. For the Old Cohort, parental education and income have been replaced with own education and income. All models include birth year and metropolitan area fixed effects. Immigrants refer to individuals born in, or that have at least one parent born in, a non-Western country. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table B.16: The reduced form effect of neighborhood composition on natives, sensitivity table

	9th grade GPA	9th grade Swedish	High school GDP	Years of schooling	Empl. Sample	Empl. Income
Panel A: Young Cohort						
Population Density	-2.087**	-1.530**	-1.233**	-0.078	-0.005	-0.081
	(0.938)	(0.748)	(0.536)	(0.050)	(0.008)	(0.100)
Excl. Neighborhoods with 100% Tipped Neighbors	-2.086*	-1.696*	-1.127*	-0.063	0.001	0.013
	(1.134)	(0.942)	(0.586)	(0.054)	(0.007)	(0.096)
Excluding Outliers	-1.905	-1.694*	-1.181*	0.037	0.001	-0.004
	(1.145)	(0.992)	(0.692)	(0.058)	(0.008)	(0.105)
Panel B: Middle Cohort						
Population Density	-1.402*	-1.912**	-0.034	-0.039	0.001	-0.005
	(0.724)	(0.828)	(0.756)	(0.056)	(0.007)	(0.088)
Excl. Neighborhoods with 100% Tipped Neighbors	-1.209	-2.001**	-0.046	-0.020	0.002	0.015
	(0.732)	(0.867)	(0.771)	(0.059)	(0.007)	(0.092)
Excluding Outliers	-0.947	-1.677*	0.390	0.009	-0.001	0.008
	(0.681)	(0.901)	(0.893)	(0.060)	(0.007)	(0.101)
Panel C: Old Cohort						
Population Density					-0.010	-0.151
					(0.012)	(0.163)
Excl. Neighborhoods with 100% Tipped Neighbors					-0.002	-0.044
					(0.012)	(0.168)
Excluding Outliers					-0.010	-0.120
					(0.012)	(0.168)

Notes: The unit of observation is an individual that resided in one of the 520 neighborhoods not used for identifying the location of the tipping points. The results are obtained from estimating equation (3). All specifications include a quartic polynomial in the difference between the neighborhood's minority share and the estimated tipping point. Standard errors are clustered on one percent bins of the running variable. Demographic controls are gender, mother's education, father's education, parental income and binaries for whether this information was not available for the individual, all measured in the base year. For the Old Cohort, parental education and income have been replaced with own education and income. All models include birth year and metropolitan area fixed effects. All models include birth year and municipality fixed effects. Natives refer to individuals not born in, and that do not have a parent born in, a non-Western country. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

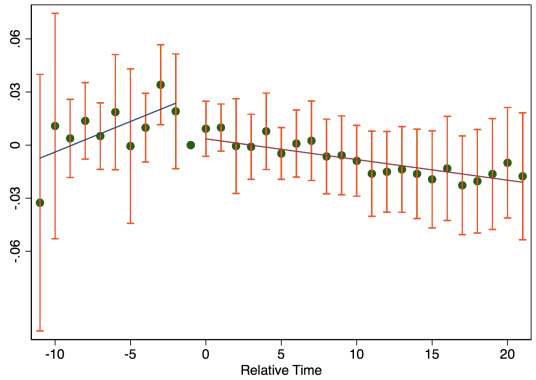
Figure C.1: Event Study Estimates by Gender and Race/Ethnicity - Hours Worked



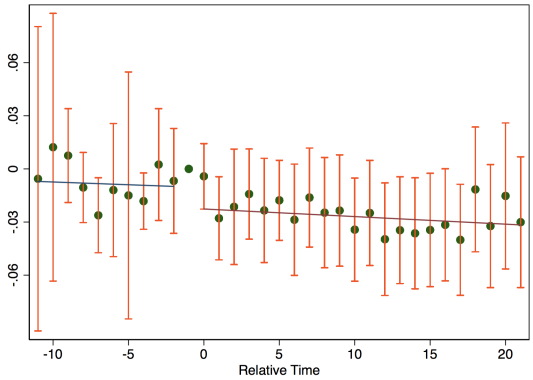
Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Relative year -1 is omitted, so all estimates are in relationship to this year. Relative year -11 includes all observations with relative time $\leq$-11 and relative year 21 includes all observations with relative time $\geq$21. All estimates include birth cohort-by-year, birth state, and year fixed effects as well as controls for exposure to school finance reforms, state EITC rates, and food stamps. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-race-gender cell. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the state level.

Figure C.2: Event Study Estimates by Gender and Race/Ethnicity - Employment

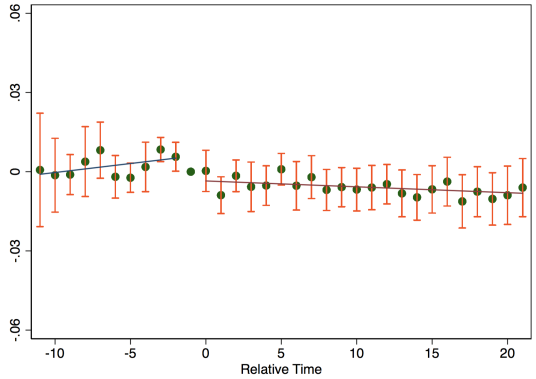
(a) Black & Hispanic Males



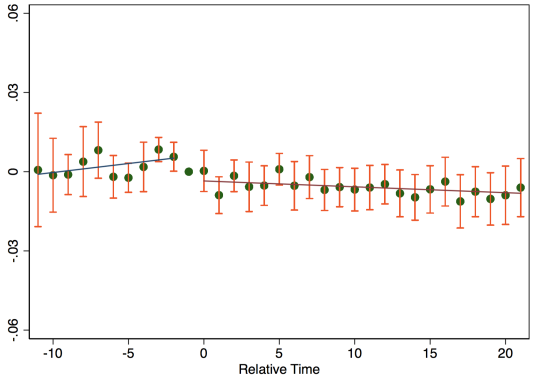
(b) Black & Hispanic Females



(c) White & Asian Males



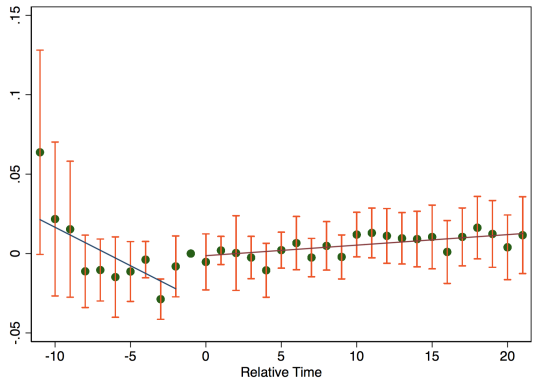
(d) White & Asian Females



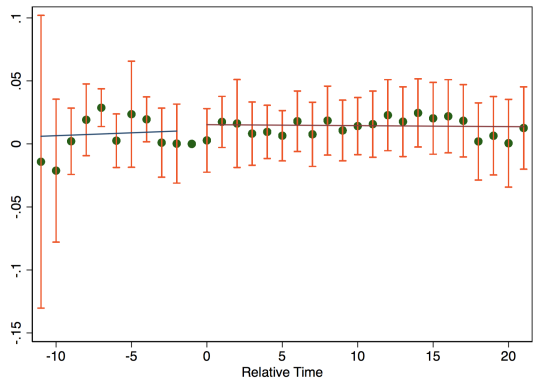
Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Relative year -1 is omitted, so all estimates are in relationship to this year. Relative year -11 includes all observations with relative time $\leq$-11 and relative year 21 includes all observations with relative time $\geq$21. All estimates include birth cohort-by-year, birth state, and year fixed effects as well as controls for exposure to school finance reforms, state EITC rates, and food stamps. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-race-gender cell. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the state level.

Figure C.3: Event Study Estimates by Gender and Race/Ethnicity - Not in Labor Force

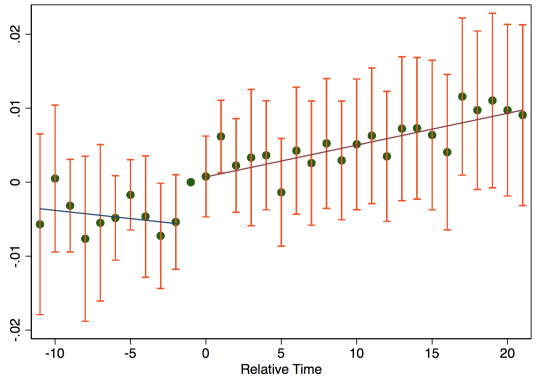
(a) Black & Hispanic Males



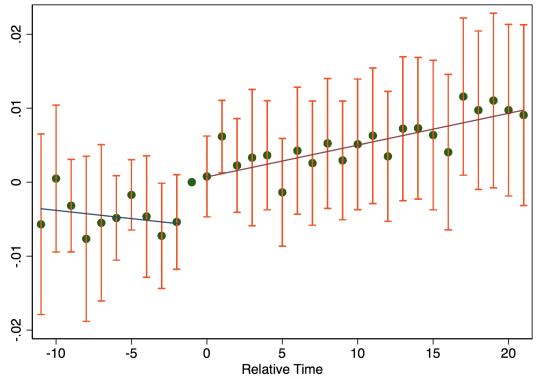
(b) Black & Hispanic Females



(c) White & Asian Males



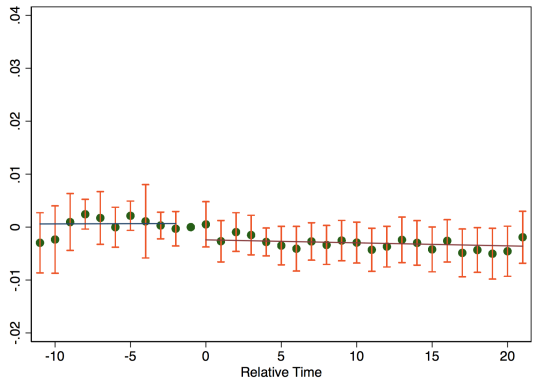
(d) White & Asian Females



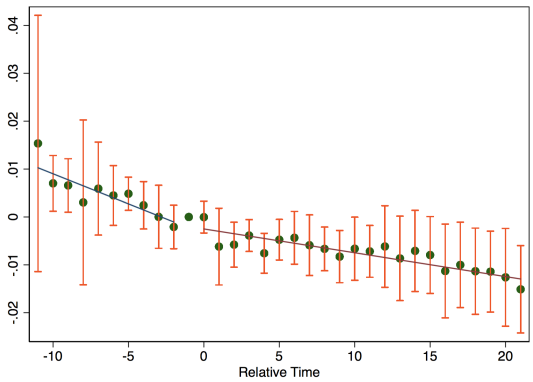
Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Relative year -1 is omitted, so all estimates are in relationship to this year. Relative year -11 includes all observations with relative time $\leq$-11 and relative year 21 includes all observations with relative time $\geq$21. All estimates include birth cohort-by-year, birth state, and year fixed effects as well as controls for exposure to school finance reforms, state EITC rates, and food stamps. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-race-gender cell. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the state level.

Figure C.4: Event Study Estimates by Gender and Race/Ethnicity - Occupational Skill

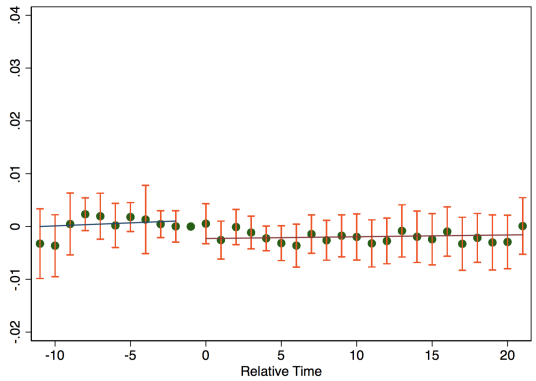
(a) Black & Hispanic Males



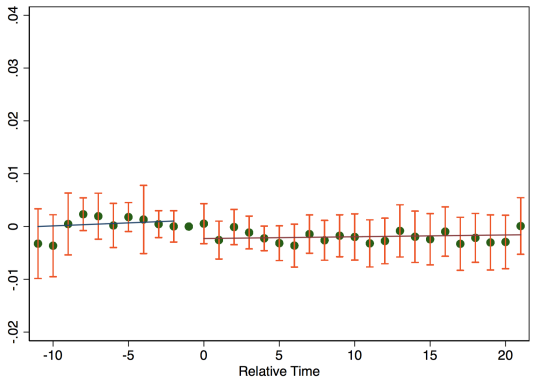
(b) Black & Hispanic Females



(c) White & Asian Males



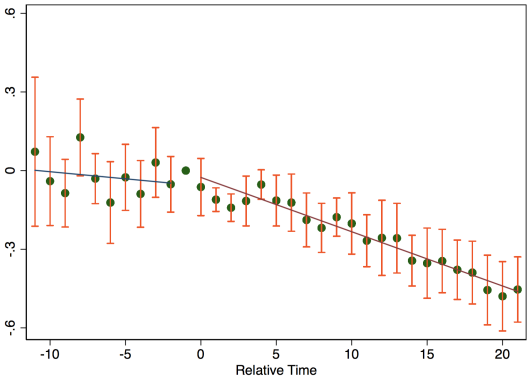
(d) White & Asian Females



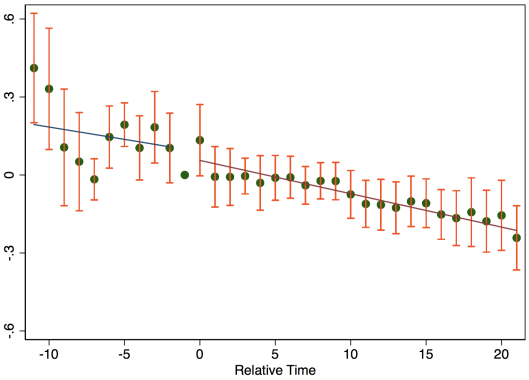
Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Relative year -1 is omitted, so all estimates are in relationship to this year. Relative year -11 includes all observations with relative time $\leq$-11 and relative year 21 includes all observations with relative time $\geq$21. All estimates include birth cohort-by-year, birth state, and year fixed effects as well as controls for exposure to school finance reforms, state EITC rates, and food stamps. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the state level.

Figure C.5: Event Study Estimates by Gender and Race/Ethnicity - Years of Education

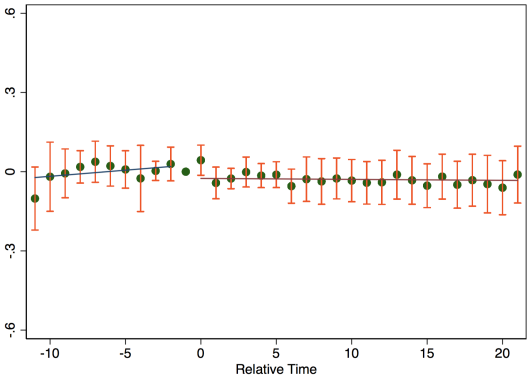
(a) Black & Hispanic Males



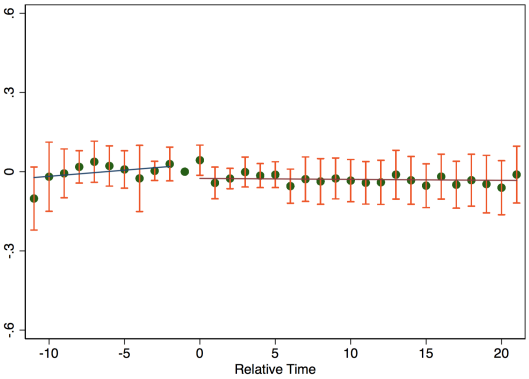
(b) Black & Hispanic Females



(c) White & Asian Males



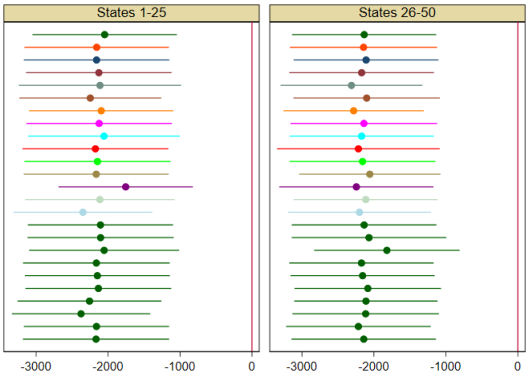
(d) White & Asian Females



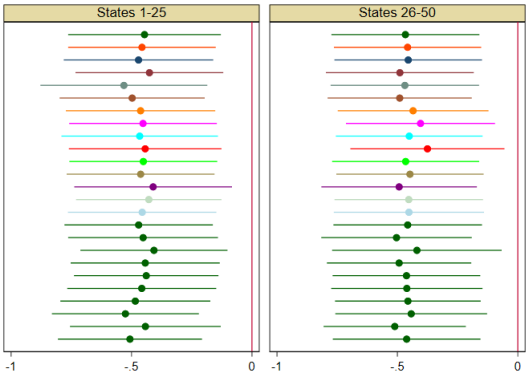
Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Relative year -1 is omitted, so all estimates are in relationship to this year. Relative year -11 includes all observations with relative time $\leq$-11 and relative year 21 includes all observations with relative time $\geq$21. All estimates include birth cohort-by-year, birth state, and year fixed effects as well as controls for exposure to school finance reforms, state EITC rates, and food stamps. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-race-gender cell. Each point is a relative time parameter estimate, while the bars extending from each point show the bounds of the 95% confidence interval calculated from standard errors that are clustered at the state level.

Figure C.6: Sensitivity of Results to Excluding Each State - Men

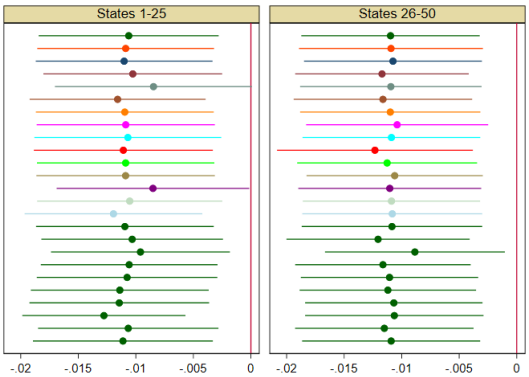
(a) Income



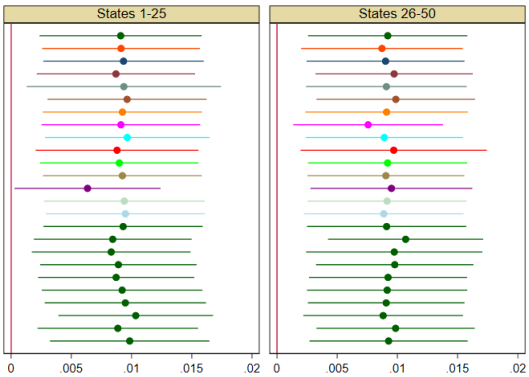
(b) Hours Worked



(c) Employment



(d) Not in Labor Force



Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Each point represents a point estimate excluding a given state from the regression and the lines extending from each point show the 95% confidence interval calculated using standard errors that are clustered at the state level.

Table C.1: Summary Statistics of Analysis Variables

	Men		Women	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Age	42.426	4.307	42.456	4.308
Asian	0.010	0.033	0.010	0.033
Black	0.128	0.096	0.144	0.106
Hispanic	0.064	0.088	0.063	0.088
Other	0.010	0.021	0.010	0.023
DTB	0.625	0.484	0.619	0.486
Years Exposed	4.710	5.613	4.646	5.599
Average EITC	0.001	0.011	0.001	0.011
Court-Ordered School Finance Reform	0.993	3.125	0.981	3.106
Legislative School Finance Reform	1.585	3.785	1.554	3.752
Food Stamp Exposure	0.625	0.325	0.621	0.326
Total Income	54,295.50	8,562.10	30,332.68	4,561.59
Hours Worked	38.964	2.112	29.552	1.685
Employed	0.822	0.046	0.730	0.043
Unemployed	0.057	0.025	0.048	0.020
Not in Labor Force	0.122	0.036	0.222	0.038
Years of Education	13.443	0.391	13.689	0.393
Occupational Skill Level	0.619	0.154	0.559	0.130
High School Degree	0.292	0.062	0.250	0.061
Some College	0.217	0.041	0.238	0.044
Associates Degree	0.081	0.023	0.109	0.026
Bachelors Degree	0.286	0.060	0.313	0.065

Notes: Authors' tabulations from 2005-2012 ACS data on 35-49 year old respondents. Tabulations are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell.

Table C.2: Summary Statistics of Analysis Variables By Gender and Race/Ethnicity

	Black & Hispanic Men		White & Asian Men		Black & Hispanic Women		White & Asian Women	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Age	42.039	4.319	42.506	4.300	42.100	4.336	42.537	4.298
Duty-to-Bargain Law	0.503	0.500	0.650	0.477	0.491	0.500	0.648	0.478
Years Exposed	3.340	5.216	4.993	5.650	3.254	5.173	4.962	5.644
Average EITC	0.000	0.007	0.001	0.012	0.000	0.006	0.001	0.012
Court-Ordered School Finance Reform	1.271	3.535	0.936	3.030	1.199	3.439	0.932	3.023
Legislative School Finance Reform	1.264	3.469	1.651	3.844	1.218	3.409	1.630	3.821
Food Stamp Exposure	0.665	0.317	0.617	0.326	0.654	0.321	0.613	0.327
Total Income	34,434.89	7,630.40	59,326.87	9,233.07	26,149.84	5,273.62	31,486.16	4,861.68
Hours Worked	33.196	3.863	40.386	1.948	29.990	3.100	29.420	1.907
Employed	0.704	0.089	0.851	0.042	0.704	0.075	0.737	0.046
Unemployed	0.085	0.052	0.049	0.024	0.074	0.045	0.041	0.019
Not in Labor Force	0.211	0.074	0.100	0.032	0.222	0.069	0.222	0.043
Occupational Skill Level	0.652	0.151	0.614	0.154	0.585	0.133	0.553	0.129
Years of Education	12.641	0.485	13.644	0.397	13.042	0.486	13.870	0.411
High School Degree	0.332	0.098	0.281	0.064	0.277	0.088	0.243	0.065
Some College	0.239	0.077	0.212	0.043	0.275	0.073	0.227	0.047
Associates Degree	0.072	0.043	0.083	0.025	0.098	0.047	0.112	0.028
Bachelors Degree	0.155	0.063	0.319	0.063	0.201	0.069	0.344	0.072

Notes: Authors' tabulations from 2005-2012 ACS data on 35-49 year old respondents. Tabulations are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell.

Table C.3: The Effect of Collective Bargaining Laws At 10 Years on Years of Education, 2008-2012 ACS Years Only

Exposure Time	All Men (i)	Black & Hispanic Men (ii)	White & Asian Men (iii)	All Women (iv)	Black & Hispanic Women (v)	White & Asian Women (vi)
At 10 Years	-0.081**	-0.150*	-0.068*	0.029	-0.106**	0.054
	(0.038)	(0.089)	(0.036)	(0.059)	(0.050)	(0.074)

Notes: Authors' estimation of equation (1) as described in the text using 2008-2012 ACS data on 35-49 year old respondents. 10-year estimates from the full event study model are shown. Regressions are based on 6,000 birth state-cohort-year observations. All estimates include birth state, year, and birth cohort-by-year fixed effects as well as controls school finance reform, EITC and food stamp measures as described in the text. Estimates in columns (i) and (iv) include controls for race/ethnicity. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. Standard errors clustered at the birth state level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.4: The Effect of DTB Laws At 10 Years After Passage for Men – Robustness Checks

Exposure Time	Earnings (i)	Hours Worked (ii)	Employed (iv)	Not in Labor Force (v)	Years of Education (v)	Occup. Skill (vi)
Panel A: Excluding States that Allow Teachers to Strike						
At 10 Years	-2144.82***	-0.543***	-0.011***	0.010**	-0.047	-0.002
	(665.46)	(0.209)	(0.004)	(0.004)	(0.042)	(0.002)
Panel B: Controlling for Total Union Membership at Age 18						
At 10 Years	-1964.58***	-0.410**	-0.009**	0.008**	-0.047	-0.003
	(621.40)	(0.182)	(0.004)	(0.004)	(0.037)	(0.002)
Panel C: Controlling for Proportion Living in Urban Areas						
At 10 Years	-1947.47***	-0.424**	-0.009**	0.008**	-0.044	-0.003
	(636.86)	(0.192)	(0.004)	(0.004)	(0.037)	(0.002)
Panel D: Controlling for Riots and Violent Protests						
At 10 Years	-2025.65***	-0.419**	-0.009**	0.008**	-0.050	-0.003
	(605.32)	(0.192)	(0.004)	(0.004)	(0.037)	(0.002)
Panel E: Controlling for Current State Fixed Effects (Individual-level Data)						
10 Years	-2384.34***	-0.427**	-0.009**	0.008**	-0.072*	-0.004
	(572.51)	(0.190)	(0.004)	(0.004)	(0.044)	(0.002)
Panel F: Including Birth State-by-Year Effects						
At 10 Years	-2045.33***	-0.593**	-0.010**	0.012**	-0.061*	-0.003
	(737.00)	(0.263)	(0.005)	(0.005)	(0.036)	(0.002)
Panel G: Including Census Region-by-Cohort Fixed Effects						
At 10 Years	-2067.11**	-0.519**	-0.010**	0.010**	-0.050	-0.003*
	(685.34)	(0.232)	(0.004)	(0.004)	(0.037)	(0.002)
Panel H: Controlling for Democratic Control of State Legislature						
10 Years	-1842.56***	-0.377**	-0.009*	0.007*	-0.050	-0.003
	(650.03)	(0.171)	(0.005)	(0.004)	(0.038)	(0.002)

Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. 10-year estimates from the full event study model are shown. All estimates include birth state, year and birth cohort-by-year fixed effects. Occupational skill results and estimates in Panel E are based on individual data and control for race/ethnicity. Other outcomes are estimated using aggregated data and control for racial/ethnic composition of the state-cohort-year-gender cell. Regressions using aggregated data are weighted by the number of individual observations that are used to calculate the averages in each state-year-cohort-gender cell. The construction of each analysis sample and control variable is described in the text. Standard errors clustered at the birth state level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.5: The Effect of DTB Laws At 10 Years After Passage for Women – Robustness Checks

Exposure Time	Earnings (i)	Hours Worked (ii)	Employed (iv)	Not in Labor Force (v)	Years of Education (v)	Occup. Skill (vi)
Panel A: Excluding States that Allow Teachers to Strike						
At 10 Years	-466.87	-0.563***	-0.014***	0.010	-0.030	-0.001
	(393.49)	(0.281)	(0.005)	(0.006)	(0.036)	(0.002)
Panel B: Controlling for Total Union Membership at Age 18						
At 10 Years	-247.82	-0.473*	-0.012**	0.008	-0.013	-0.002
	(343.85)	(0.274)	(0.005)	(0.006)	(0.033)	(0.002)
Panel C: Controlling for Proportion Living in Urban Areas						
At 10 Years	-245.41	-0.469*	-0.011**	0.008	-0.012	-0.002
	(360.22)	(0.262)	(0.005)	(0.006)	(0.032)	(0.002)
Panel D: Controlling for Riots and Violent Protests						
At 10 Years	-272.60	-0.470*	-0.012**	0.008	-0.013	-0.002
	(369.72)	(0.269)	(0.005)	(0.006)	(0.033)	(0.002)
Panel E: Controlling for Current State Fixed Effects (Individual-level Data)						
At 10 Years	-399.70	-0.453*	0.003	0.009	-0.031	-0.001
	(337.58)	(0.252)	(0.003)	(0.006)	(0.033)	(0.002)
Panel F: Including Birth State-by-Year Effects						
At 10 Years	-260.61	-0.482	-0.011*	0.010	-0.011	-0.002
	(390.95)	(0.314)	(0.006)	(0.007)	(0.036)	(0.002)
Panel G: Including Census Region-by-Cohort Fixed Effects						
At 10 Years	-312.31	-0.495*	-0.012**	0.008	-0.010	-0.002
	(379.21)	(0.282)	(0.005)	(0.007)	(0.033)	(0.002)
Panel H: Controlling for Democratic Control of State Legislature						
At 10 Years	-363.04	-0.456*	-0.013***	0.008	-0.018	-0.002
	(360.69)	(0.272)	(0.005)	(0.006)	(0.033)	(0.002)

Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. 10-year estimates from the full event study model are shown. All estimates include birth state, year and birth cohort-by-year fixed effects. Occupational skill results and estimates in Panel E are based on individual data and control for race/ethnicity. Other outcomes are estimated using aggregated data and control for racial/ethnic composition of the state-cohort-year-gender cell. Regressions using aggregated data are weighted by the number of individual observations that are used to calculate the averages in each state-year-cohort-gender cell. The construction of each analysis sample and control variable is described in the text. Standard errors clustered at the birth state level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.6: The Correlation of Duty-to-Bargain Exposure with Fixed Individual Characteristics and State Observables Unrelated to Collective Bargaining

	Panel A: Men						
	Age (x100)	Black	Hispanic	Asian	Other Race	Fraction Homeowner	Fraction State Male
	(i)	(ii)	(iv)	(v)	(v)	(vi)	(vii)
Relative Years to DTB	-0.000	0.002**	-0.0004	0.0003***	-0.0000	0.0002	-0.0010***
	(0.000)	(0.001)	(0.0013)	(0.0001)	(0.0001)	(0.0006)	(0.0003)
I(DTB)	-0.000*	0.009**	0.002	-0.001	-0.001	0.002	0.003***
	(0.000)	(0.004)	(0.007)	(0.001)	(0.001)	(0.002)	(0.001)
Relative Years to DTB)	-0.000*	-0.0001	0.003	0.0000	0.0001	-0.0003	0.0007**
*I(DTB)	(0.000)	(0.0009)	(0.002)	(0.0001)	(0.0001)	(0.0006)	(0.0003)
	Panel B: Women						
	Age	Black	Hispanic	Asian	Other Race	Fraction Homeowner	Fraction State Male
	(i)	(ii)	(iv)	(v)	(v)	(vi)	(vii)
Relative Years to DTB	-0.000	0.003*	-0.001	0.0002***	0.0000	0.0009***	0.0001
	(0.000)	(0.002)	(0.0013)	(0.001)	(0.0001)	(0.0003)	(0.0003)
I(DTB)	-0.000**	0.005	0.003	0.001	-0.0012**	-0.004***	-0.003
	(0.000)	(0.005)	(0.006)	(0.001)	(0.0007)	(0.001)	(0.003)
Relative Years to DTB)	-0.000	-0.001	0.003	0.0001	0.0001	-0.0007**	-0.0001
*I(DTB)	(0.000)	(0.002)	(0.002)	(0.0001)	(0.0001)	(0.0003)	(0.0003)

Notes: Authors' estimation of equation (2) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. Relative Years to DTB Law is the number of years relative to the passage of a duty-to-bargain law when each cohort was 6 years old, which is set to zero for states that never pass such a law. I(DTB Law) is an indicator for whether a duty-to-bargain law has been passed in the state when each cohort was 6 years old. All estimates include state, year and birth cohort-by-year fixed effects. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. State-specific outcomes are averaged over the individual ACS observations, which is why the male and female estimates differ numerically for these outcomes. Standard errors clustered at the birth state level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.7: The Effect of Collective Bargaining Laws At 10 Years on Long-Run Outcomes for Men – Accounting for Mobility

Panel A: Dropping Those Who do not Live in State of Birth						
Exposure Time	Earnings (i)	Hours Worked (ii)	Employed (iv)	Not in Labor Force (v)	Years of Education (v)	Occup. Skill (vi)
At 10 Years	-2161.72**	-0.611**	-0.006	0.010**	-0.079	-0.005**
	(1051.35)	(0.307)	(0.006)	(0.005)	(0.050)	(0.002)
Panel B: Weighting by Childhood Mobility						
Exposure Time	Earnings (i)	Hours Worked (ii)	Employed (iv)	Not in Labor Force (v)	Years of Education (v)	Occup. Skill (vi)
At 10 Years	-1831.58***	-0.511**	-0.004	0.003	-0.041	
	(532.85)	(0.110)	(0.003)	(0.003)	(0.033)	

Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. 10-year estimates from the full event study model are shown. All estimates include state, year and birth cohort-by-year fixed effects, as well as controls for the racial/ethnic composition of the state-cohort-year-gender cell, exposure to school finance reforms, average state EITC and average food stamp availability during school years. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. In Panel (A), we exclude the 37.7% of respondents who do not live in their state of birth. In Panel (B), we expand the data to be at the state of birth-cohort-potential migration state level and weight each observation by the proportion of 17 year olds in the 1990 census who were born in the birth state and lived in the migration state. All variables are defined using the migration state. Standard errors clustered at the birth state level in Panel (A) and two-way clustered at the birth state and migration state in Panel (B) are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.8: The Effect of Collective Bargaining Laws At 10 Years on Long-Run Outcomes for Women – Accounting for Mobility

Panel A: Dropping Those Who do not Live in State of Birth						
Exposure Time	Earnings (i)	Hours Worked (ii)	Employed (iv)	Not in Labor Force (v)	Years of Education (v)	Occup. Skill (vi)
At 10 Years	-199.11	-0.406	-0.013**	0.010	-0.004	0.000
	(419.60)	(0.317)	(0.006)	(0.007)	(0.049)	(0.003)
Panel B: Weighting by Childhood Mobility						
Exposure Time	Earnings (i)	Hours Worked (ii)	Employed (iv)	Not in Labor Force (v)	Years of Education (v)	Occup. Skill (vi)
At 10 Years	-477.78*	-0.248	-0.005	0.003	-0.018	
	(273.55)	(0.198)	(0.004)	(0.005)	(0.025)	

Notes: Authors' estimation of equation (1) as described in the text using 2005-2012 ACS data on 35-49 year old respondents. 10-year estimates from the full event study model are shown. All estimates include state, year and birth cohort-by-year fixed effects, as well as controls for the racial/ethnic composition of the state-cohort-year-gender cell, exposure to school finance reforms, average state EITC and average food stamp availability during school years. Regressions are weighted by the number of individual observations that are used to calculate the averages in each state-cohort-year-gender cell. In Panel (A), we exclude the 37.7% of respondents who do not live in their state of birth. In Panel (B), we expand the data to be at the state of birth-cohort-potential migration state level and weight each observation by the proportion of 17 year olds in the 1990 census who were born in the birth state and lived in the migration state. All variables are defined using the migration state. Standard errors clustered at the birth state level in Panel (A) and two-way clustered at the birth state and migration state in Panel (B) are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.9: The Relationship Between Duty-to-Bargain Laws and School Resources

	Dependent Variable: Log of				
Independent Variable	Teacher Salary Expenditures (i)	Administrative Salary Expenditures (ii)	Other Salary Expenditures (iv)	Teacher-Student Ratio (v)	Operating Expenditures per Student (v)
Relative Years to DTB	-0.038***	-0.064	0.053***	-0.005	-0.005
	(0.011)	(0.038)	(0.012)	(0.003)	(0.020)
I(DTB)	0.080	0.470*	-0.196***	0.044	0.057
	(0.066)	(0.269)	(0.066)	(0.031)	(0.059)
(Relative Years to DTB)	0.038***	0.089**	-0.044***	-0.005	0.008
*I(DTB)	(0.010)	(0.022)	(0.007)	(0.009)	(0.009)

Notes: Authors' estimation of equation (2) as described in the text using 1972-1991 Census/Survey of Governments Data. The data vary at the state-year level and all estimates include state and year fixed effects. Regressions are weighted by total enrollment in each state. Relative Years to DTB Law is the number of years relative to the passage of a duty-to-bargain law when each cohort was 6 years old, which is set to zero for states that never pass such a law. I(DTB Law) is an indicator for whether a duty-to-bargain law has been passed in the state when each cohort was 6 years old. All outcome variables are in logs, and salary expenditures reflect total expenditures on each category including part-time and full-time teachers. Standard errors clustered at the state level are in parentheses: *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.