THREE ESSAYS IN APPLIED MICROECONOMICS

By

Tucker Weldon Smith

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

May 10th, 2024

Nashville, Tennessee

Approved:

Lesley Turner, Ph.D.

Brian Beach, Ph.D.

Michelle Marcus, Ph.D.

Christopher Candelaria, Ph.D.

To Mom, Dad, Cole, Logan, Gigi, Ma, and Nanny.

ACKNOWLEDGMENTS

For her constant guidance and support, I thank Lesley Turner. I thank my dissertation committee members Brian Beach, Michelle Marcus, and Christopher Candelaria for their feedback and support. I also thank Riley Acton, Adam Blandin, Eric Bond, Kitt Carpenter, Bill Collins, Mitchell Downey, Patrick Flynn, Andrew Foote, Andrew Goodman-Bacon, Nora Gordon, Michel Grosz, Vikram Maheshri, Daniel Mangrum, Camila Morales, Orgul Ozturk, Analisa Packham, Matthew Pesner, Joel Rodrigue, Amanda Ross, Peter Schott, Bryan Stuart, Christopher Timmins, and Patrick Turner for their valuable insights and support of my research. I am grateful for feedback from seminar participatants at the Federal Reserve Bank of Dallas, University of South Carolina, University of Alabama, University of North Carolina at Charlotte, Western Kentucky University, Vanderbilt Universtity, the Southern Economic Association Annual Meeting, the Association for Education Finance and Policy Annual Conference, Urban Economics Association PhD Student Workshop, and the Urban Economics Association North American Meeting. This dissertation was supported by a grant from the American Educational Research Association which receives funds for its“AERA-NSF Grants Program” from the National Science Foundation under NSF award NSF-DRL 1749275. Opinions reflect those of the author and do not necessarily reflect those AERA or NSF. This dissertation was also supported by the Walter B. Noel Dissertation Fellowship, the Kirk Dornbush Summer Research Grant, and the College of Arts and Science Summer Research Award. The conclusions of this research also do not necessarily reflect the opinion or official position of the University of Houston Education Research Center, the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas. This dissertation would not have been completed without the outstanding undergraduate teaching and advising from Liz Bouzarth, Nathan Cook, John Harris, Kevin Hutson, Jason Jones, Kailash Khandke, Ken Peterson, Tom Smythe, and Jeffrey Yankow at Furman University. Finally, I am grateful for the support and encouragement from my parents, Allen and Debbie; my brother, Cole; my grandmothers, Anne and Nancy; my friends, Nate Boehm, Luis Carvajal-Osorio, Sam Crowell, Kaitlyn Elgart, Tucker Erdmann, Rowan Isaaks, Jessica Kiser, Elliot Loftus, Frankie Wiley, and Eric Williams; and the faculty in the Vanderbilt University Department of Economics.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Human Capital Adjustments and Labor Market Resilience: Evidence from Linked Education and Earnings Data

1.1 Introduction

Negative labor demand shocks can cause remarkably persistent harmful effects on exposed individuals. Both established workers exposed to a labor demand shock and individuals first entering the labor market during periods of depressed labor demand experience sustained declines in earnings, known as "scarring effects" (e.g., Ruhm, 1991; Kahn, 2010). Whether individuals coming of age can avoid the same fate likely depends on their ability to make extensive-margin (i.e., attending college) and intensive-margin (i.e., field of study) adjustments to their human capital investments. As automation, decarbonization, and other looming changes to the U.S. economy threaten to bear uneven benefits and costs across workers and local labor markets, understanding the degree to which human capital adjustments can protect against persistent earnings losses is particularly important.

In this paper, I study the effects of exposure during youth and adolescence to negative local labor demand shocks generated by Chinese import competition (i.e., the "China shock") on human capital accumulation and later-life earnings. Using linked student-level administrative data from Texas, I find that students from counties exposed to larger local shocks were 4% more likely to enroll in college and 8% more likely to earn a bachelor's degree. I provide evidence that these adjustments, along with shifts of fields of study away from those directly exposed to import competition in both high school and college, shielded students from 90% of the decline in earnings experienced by young adults that had already made key educational decisions prior to the onset of the shock. These results suggest a silver lining to the gloomy findings of prior work on the long-term effects of the China shock (e.g., Autor et al., 2021) and other negative labor demand shocks (e.g., Stuart, 2022): if individuals coming of age sufficiently adjust their human capital investments, they can emerge relatively unscathed.

My research design exploits quasi-random variation in exposure to changes in local labor demand based on a change in U.S. trade policy in October 2000 – formally, the establishment of "Permanent Normal Trade Relations" (PNTR) with China. PNTR exposed subsets of domestic manufacturing firms to increased competition from Chinese exporters (Pierce and Schott, 2016). This unexpected policy change resulted in larger negative labor demand shocks in counties with more firms specializing in exposed subsets of manufactured goods (e.g., toys and games) compared to those with similar levels of manufacturing employment, but with firms producing less-exposed product specialties (e.g., processed foods) (Pierce and Schott, 2020; Greenland et al., 2019).

Building off of Pierce and Schott (2016, 2020), I use this variation in a two-step "de-trended" difference-in-differences specification (Goodman-Bacon, 2021c). I compare changes in educational and labor market outcomes across cohorts of students that reached critical ages before and after the policy change (first difference) in counties that were more- and less-exposed to PNTR (second difference) relative to existing linear trends. To account for bias from endogenous migration in response to the shock, I estimate intent-to-treat models that assign treatment to students based on the county where they attended school prior to the policy change. My preferred specification controls for individual-level demographics and allows for baseline county characteristics to flexibly affect outcomes across cohorts, but my main results are robust to excluding covariates or controlling for exposure to subsequent labor demand shocks. Identification requires the assumption that differences in outcomes of students from counties that experience large and small shocks would continue to evolve along existing differential linear trends if both groups were exposed to small labor demand shocks (Callaway et al., 2021; Goodman-Bacon, 2021c).[1] I provide support for causal interpretation by showing that differences in *fixed* student characteristics across counties that later experienced larger and smaller shocks continued along existing trends across cohorts after the policy change.[2] Moreover, I show that my main results are generally robust to relaxing this assumption by allowing trends in potential outcomes to revert to parallel after the onset of the shock and by allowing smooth deviations in potential outcomes from existing linear trends (Rambachan and Roth, 2023).

To inform interpretation of my main estimations, I first confirm that Chinese import competition had similar effects on local labor markets in Texas as previous research finds nationwide.[3] Counties with above-median exposure to PNTR experienced a 3.0 percentage-point decline in their employment-to-population ratio relative to those with below-median exposure. On average, more-exposed counties saw an 8% decline in earnings for workers without a college degree and an 18% decline in earnings for workers between the ages of 15 and 24. These substantial declines in the opportunity costs of schooling, along with increases in the college earnings premium, likely incentivized marginal students to pursue a postsecondary education instead of entering the labor force. On the other hand, declines in family income may have inhibited their ability to do so. I find that greater exposure to the shock caused an 8% increase in student eligibility for free-or-

[1]This essentially combines two assumptions: (1) outcomes of students from counties that later experience larger and smaller shocks would continue along existing differential linear trends in the absence of PNTR and (2) students from both groups of counties would respond similarly on average to a similarly sized labor demand shock (Callaway et al., 2021).

[2]This exercise is analogous to standard difference-in-differences placebo analyses (i.e., putting outcomes measured prior to treatment on the left-hand side) but adapted to my two-step specification. An alternative interpretation of the exercise is that it tests for selection on observables into treatment.

[3]Previous research on the local labor market effects of trade liberalization with China finds that more exposed localities experienced persistent decreases in employment and earnings, particularly among low-skilled workers and workers in exposed manufacturing subsectors (Autor et al., 2013, 2015; Pierce and Schott, 2016; Greenland and Lopresti, 2016; Autor et al., 2021).

reduced-price lunch, a proxy for low-income status. These effects are similar in size to estimates of earnings and employment losses across the country caused by the same shock and to those caused by a recession on exposed local labor markets.[4]

In my primary analyses, I first examine whether students made adjustments to their human capital investments in high school. Although I find that local shocks did not affect the likelihood students graduated high school – consistent with prior estimates using school-level graduation counts (Burga and Turner, 2022), I provide evidence of substantial intensive-margin human capital adjustments by forward-looking students along novel margins.[5] In high school, exposed students reduced their enrollment in manufacturing-aligned vocational elective courses and took more Advanced Placement, International Baccalaureate, and college-credit courses. These responses suggest that students internalized both reductions in long-term earnings prospects in industries directly exposed to the labor demand shock and increases in the college earnings premium.

Following students beyond high school, I estimate that greater exposure to local shocks caused a 1.8 percentage-points (4%) increase in the likelihood of enrolling in a public college in Texas.[6] This magnitude is comparable to previously estimated effect sizes of smaller elementary school classrooms (Chetty et al., 2011) or a $1,000 increase in need-based financial aid (Castleman and Long, 2016).[7] By age 25, exposed students accumulated 5% more cumulative semesters and were 1.1 percentage-points (8%) more likely to have earned a bachelor's degree. Just as in high school, students adjusted their choice of fields of study away from those directly exposed to import competition. Using the distribution of recent graduates of specific majors across industries of employment prior to the onset of PNTR to define *major-level* exposure to the policy change, I find that community college enrollment significantly increased in the least-exposed fields but fell in those most exposed to the shock.

Evidence of the persistence of the China shock's negative effects on earnings of prime-aged workers (Autor et al., 2021) and of the "scarring" effects of entering the labor market during economic downturns (e.g., Kahn, 2010) raise the question of whether the above human capital adjustments translated into improved

[4] Autor et al. (2021) finds that the China shock reduced overall employment relative to population by nearly 2 percentage-points in exposed local labor markets, and Greenland and Lopresti (2016) find the shock decreased earnings for workers without a college degree by 6%. Moreover, Hershbein and Stuart (2023) show that each recession since the 1970s corresponded to approximately a 3 percentage-point employment decline in counties with above-median exposure.

[5] Greenland and Lopresti (2016) and Burga and Turner (2022) both examine the effects of the China shock on high school graduation rates using aggregated graduation counts but come to different conclusions. Greenland and Lopresti (2016) find that graduation rates increase by 3.6 percentage-points in local labor markets exposed to the China shock; however, Burga and Turner (2022) provide evidence that this result is mostly explained by outmigration and weak instrument bias.

[6] I find that exposed students were 1.2 percentage-points more likely to enroll in a two-year Texas public community or technical college and 1.6 percentage-points more likely to enroll in a Texas four-year public university by age 20. I do not observe if a student enrolls in a college outside of Texas, and the interpretation of my estimates as representing extensive effects on college enrollment would be threatened by substitution from out-of-state or private colleges into in-state public institutions. I provide evidence supporting that such substitution did not occur in Section 1.7.

[7] Chetty et al. (2011) find that assignment to small classes (averaging 15 students, rather than 22) for grades K-3 through the Tennessee STAR experiment caused a 1.8 percentage-point increase in college enrollment. Castleman and Long (2016) find that a $1,995 (2020$) increase in need-based aid in Florida caused a 3.2 percentage-point increase in enrollment at public four-year universities. Assuming linearity, my estimated effect on four-year enrollment of 1.6 percentage-points scaled by their coefficient corresponds to a $998 increase in aid.

labor market outcomes for young adults exposed to the shock.[8] Consistent with scarring effects, individuals from high-exposure counties that were old enough to make key human capital investments before 2000 earned 8% less than those from low-exposure counties after the onset of the shock. However, exposed students young enough to adjust their educational decisions experienced statistically significant relative earnings gains large enough to erase 90% of this gap. This suggests that human capital adjustments nearly fully buffered against scarring effects of the shock, despite the persistence of the shock's negative effects on overall per-capita earnings and employment rates in exposed local labor markets.

Estimates of the average effects of the shock on human capital accumulation and later-life earnings may mask substantial scarring of "left-behind" students. Therefore, I examine heterogeneous responses by particularly vulnerable subgroups of students. I find that male students, students from low-income households, and students identifying as racial or ethnic minorities all adjusted their human capital investments in manners that yielded substantial protection against the labor demand shock.

Previous work finds mixed evidence on the effects of exposure to negative labor demand shocks during youth and adolescence on human capital accumulation.[9] My rich administrative dataset allows me to make two contributions to this literature, both of which inform the interpretation of previous findings. First, I find evidence of key adjustments along novel dimensions such as high-school course-taking that suggest forward-looking students observed and responded to salient changes in the returns to education across fields of study and attainment levels. Students may also have made similar adjustments in other settings where previous research finds limited responses on "extensive" margins of educational attainment (Stuart, 2022; Burga and Turner, 2022; Ferriere et al., 2018). Thus, "null" results in these settings may still reflect meaningful changes in human capital accumulation. Moreover, changes in expectations and forward-looking adjustments in high school plausibly aided complementary shifts in fields of study in college, potentially explaining why I find "larger" shifts of fields than prior work examining effects of salient labor demand shocks occurring closer to college entrance (Acton, 2021).[10]

Second, I provide evidence that human capital adjustments in my setting translated into labor market

[8] Kahn (2010) finds that a one percentage-point increase in the state unemployment rate at the time of labor market entry decreases the wages of college graduates by 9%, an effect that persists at least 15 years after college graduation. Oreopoulos et al. (2012) and Altonji et al. (2016) similarly find persistent declines in wages for college graduates entering the labor markets during a recession, and Hershbein (2012) shows smaller (1-2%) and less persistent wage declines for high school graduates. Schwandt and von Wachter (2019) finds that entering the labor market when your birth state exhibits a high unemployment rate has substantial negative effects on earnings that persist for 10 years and are strongest for non-white workers and high school dropouts.

[9] Work in this area analyzes local shocks caused by recessions (Stuart, 2022; Weinstein, 2022), mass layoffs (Foote and Grosz, 2020; Acton, 2021; Salvanes et al., 2021), automation (Di Giacomo and Lerch, 2021), housing busts (Charles et al., 2018), energy busts(Black et al., 2005), deindustrialization (Choi, 2023), and trade (Greenland and Lopresti, 2016; Ferriere et al., 2018; Lee, 2021; Burga and Turner, 2022).

[10] Acton (2021) finds that high school graduates responded to local plant closures in their senior year of high school by shifting community college program choices from those related to jobs lost in the closure and toward occupations that required similar skillsets. Examining a longer time horizon and a more "permanent" shock, I find that students exposed to the China shock before entering high school were less likely to enroll in manufacturing-aligned community college programs and more likely to enroll in health, IT, or business programs.

benefits. In doing so, I bridge the above literature with that on the scarring effects of negative labor demand shocks on earnings. Stuart (2022) finds that exposure to the 1980-1982 recession during childhood or adolescence caused decreases in college degree receipt and long-term earnings and attributes these effects to decreases in "childhood investments" in human capital. In contrast, I find that students exposed to the China shock before entering high school – and as early as in kindergarten – adjusted their educational decisions in manners that protected them against similar scarring effects on earnings. I show that Texas' "Robin Hood" K-12 finance system prevented declines in local property values from manifesting in declines in school spending – a key component of "childhood" human capital investments, potentially contributing to the differences in our results.[11]

The literature on scarring effects also finds that individuals that enter a labor market during a period of depressed labor demand experience lasting declines in earnings (Kahn, 2010; Hershbein, 2012; Oreopoulos et al., 2012; Schwandt and von Wachter, 2019). One proposed explanation for these scarring effects is that individuals entering the labor market during downturns disproportionately possess skills that do not match those demanded by employers (Liu et al., 2016). Consistent with this hypothesis, I find evidence suggesting that adjustments of human capital investments to align with changes in local labor demand across sectors dramatically reduce scarring effects on average, despite the persistence of overall declines in local labor demand and earnings losses for prime-aged workers.

Finally, this paper contributes to existing research on the China shock. In contrast to classical trade models, this literature finds substantial and persistent consequences of exposure to Chinese import competition on exposed local labor markets and workers (e.g., Autor et al., 2014, 2021). I provide evidence that, at least in Texas, individuals coming of age largely avoided such harm, and my results suggest that the extent to which the costs of other skill-biased technological changes (e.g., automation) transmit across generations may also depend on the ability of students to adjust their human capital investments and the presence of policies to support these investments.[12]

1.2 Background

My research design leverages variation in declines in local labor demand resulting from a change in U.S. trade policy toward China in 2000. The U.S. subjects goods imported from foreign countries to one of two sets of tariff rates. Goods imported from fellow members of the World Trade Organization (WTO) are subject to relatively low "column 1" rates (hereafter, "preferred" tariff rates), while goods from nonmarket economies are subject to relatively high "column 2" rates (hereafter, "punitive" tariff rates) set by the Smoot-Hawley

[11] Other differences in my empirical setting and that of (Stuart, 2022) may contribute to differences in our respective estimates of the effects of exposure to negative labor demand shocks on college degree receipt and later-life earnings. These include the greater presence of need-based financial aid programs in my sample period and the *ex ante* permanence of the China shock, among others.

[12] I discuss what such policies might resemble in Section 1.8

Tariff Act of 1930. In 1999, preferred tariff rates averaged 4% and punitive rates averaged 37%. Moreover, the difference between the preferred and punitive tariff rates (the "tariff gap") varied widely by type of good, ranging from a 0 to 80 percentage-point difference.

The President may annually extend preferred tariff rates to nonmarket economies, although Congress can pass legislation to block such an extension. The U.S. first granted preferred tariff rates to Chinese imports in 1980; however, following the Tiananmen Square incident in 1989, Congressional approval of these preferred tariff rates became a politically contentious process.[13] The political uncertainty surrounding trade policy with China permeated into the operations of U.S. firms, and those with large tariff gaps were particularly unlikely to outsource production to China due to the prospect of reversal to punitive tariff rates in any given year (Pierce and Schott, 2016).[14]

In October 2000, Congress passed a bill to establish "Permanent Normal Trade Relations" (PNTR) with China, permanently locking in preferred tariff rates for Chinese goods imported into the U.S., and the following December, China formally joined the World Trade Organization (WTO). In response to the differential change in incentives, the real value of U.S. imports from China increased for goods with high tariff gaps, such as toys, relative to those with low tariff gaps, such as processed foods (Pierce and Schott, 2016). Intuitively, areas with larger shares of workers in industries with high tariff gaps experienced larger reductions in local labor demand due to import competition following the policy change: Pierce and Schott (2020) find that an interquartile shift in a county's tariff gap was associated with approximately a 1 percentage-point increase in the unemployment rate and a 1.5 percentage-point decline in labor force participation by 2007. Overall, the real value of Chinese imports nearly tripled by 2007; during this same period, U.S. domestic manufacturing employment fell by over 3 million workers.

Research on the China shock finds that exposed local labor markets experienced increases in plant closures and sharp and sustained reductions in employment, labor force participation, and income (Acemoglu et al., 2014; Autor et al., 2013, 2021). Employment and earnings losses were concentrated among low-skilled and less-educated workers (Autor et al., 2013, 2014; Acemoglu et al., 2014; Pierce and Schott, 2016), and the existing literature provides mixed evidence as to whether such effects spilled over into industries outside

[13]Legislation to block preferred tariff rates from being applied to Chinese goods was voted on in the U.S. House of Representatives annually from 1990 until 2001 and passed the House in 1990, 1991, and 1992, but never passed the Senate.

[14]The chilling effect of uncertainty over tariff rates on U.S. firms is evident by a 1993 letter from CEOs of 340 U.S. firms to President Clinton describing the annual tariff renewal process as creating "an unstable and excessively risky environment fo US companies considering trade and investment in China" (Rowley, 1993)

of manufacturing.[15] These effects persisted through at least 2016 (Autor et al., 2021). Additionally, more exposed labor markets experienced reductions in housing prices and public spending (Feler and Senses, 2017) and increases in fatal drug overdoses (Pierce and Schott, 2020; Autor et al., 2019).

The above literature on the local labor market effects of the China shock suggest multiple channels through which exposure to the could would affect human capital accumulation according to canonical human capital theory (Becker, 1962). First, diminished *contemporaneous* labor market opportunities for school-aged workers represent declines to the opportunity costs of schooling, incentivizing the extensive-margin adjustments of high school completion and college attendance. Second, the shock's ex ante "permanence" and incidence on workers with low levels of education and those in the manufacturing sector plausibly shifted expected *lifetime* earnings across education levels and fields of study. Such changes to earnings premia may cause both extensive-margin and intensive-margin changes to human capital investments. Decreased expected earnings for high school dropouts and high school graduates relative to those with a college education would increase the expected college earnings premium, further incentivizing college enrollment and – to the extent that students are forward-looking – intermediate changes to high school course-loads that support college preparation. Concurrently, decreases in expected earnings for workers in sectors directly exposed to the labor demand shocks, such as manufacturing, would reduce the expected returns for studying fields that lead to employment in these industries and incentivize students to adjust their course and major selection away from such fields.

Finally, the negative shock to local labor demand may tighten constraints on "external" investments in human capital by a student's family and school.[16] The specifics of my setting allow me to examine the effects of a negative shock to local labor demand while holding constant – in relative terms – K-12 (kindergarten through 12th-grade) school spending.[17] Since 1993, Texas has employed a relatively unique school finance equalization system known as the "Robin Hood" plan. Under order by the Texas Supreme Court, the state education agency equalizes spending across districts by redistributing excess funding from property-wealthy districts toward property-poor districts.[18] As a result of this policy, even though K-12 education in Texas

[15] Using a shift-share design and aggregate Census data, Autor et al. (2013) find that local labor markets (defined as commuting zones) exposed to the China shock experienced decreases in non-manufacturing employment relative to population, and this reduction statistically differs from zero for workers without a college education. Bloom et al. (2019) utilize the same design and establishment-level data from the Census Bureau's Longitudinal Business Database and find commuting-zone-level exposure to the China shock resulted in small increases in non-manufacturing employment. Both papers find reductions in manufacturing and overall employment in exposed local labor markets. Moreover, Ahlquist and Downey (2023) find that spouses and children of manufacturing workers exposed to the China shock were more likely to find work in education, social work, and health care. Pierce et al. (2022) show that workers employed outside of manufacturing even experienced earnings gains in counties with large clusters of employment "downstream" of exposed manufacturing industries, due to reductions in input costs.

[16] Both theoretical and empirical research provides evidence that investments in human capital during childhood can yield substantial gains in later-life skills and earnings through dynamic complementarities (Heckman and Mosso, 2014; Almond et al., 2018; Johnson and Jackson, 2019; Stuart, 2022)

[17] See Jackson (2018) for a review of the recent literature on the causal effects of school spending on student outcomes.

[18] The Robin Hood school finance equalization system was originally adopted in response to the Texas Supreme Court case *Edgewood Independent School District v. Kirby*. Appendix 1.9.3.1 details the Robin Hood system.

is largely funded through property taxes, school spending is nearly orthogonal to changes in *local* property wealth. I show in Section 1.4.2 that the Robin Hood formula shielded students in counties exposed to adverse local shocks from reductions in school spending that would typically accompany declines in local property wealth. However, I also provide evidence that students from more-exposed counties experienced declines in family income. This may particularly hinder students from adjusting by attending college, which often requires a substantial monetary cost.[19]

Nearly 11% of all kindergarten through 12th-grade (K-12) students and 10% of full-time college students in the U.S. attend school in Texas, making the state an inherently relevant setting to study the effects of labor demand shocks on human capital accumulation. Public K-12 schools in Texas differed from those of the U.S. as a whole in the years before the trade shock in terms of student demographics – 40% of students in public K-12 schools in Texas were Hispanic, as opposed to only 16% across the country overall – but received similar per-pupil funding levels and produced similar high school graduation rates.[20]

1.3 Data

My primary dataset consists of administrative student-level data from the University of Houston Education Research Center (UHERC), featuring individual-level linked K-12, postsecondary, and workforce records from 1994 to present. K-12 data from the Texas Education Agency include standardized test scores, course schedules, graduation records, and demographic information for all students who attended public schools in Texas. Postsecondary data from the Texas Higher Education Coordinating Board include enrollment, declared majors, and degree and certificate records from all public colleges and universities in the state. Finally, wage reports from the Texas Workforce Commission include monthly earnings and employment records by employer for all Texas workers earning wages in positions covered by state unemployment insurance.[21]

To build my analysis sample, I start with cohorts of students that entered ninth grade at Texas public high schools since Fall 1995. Two factors related to the nature of the administrative data determine further selection. First, I only observe students attending public schools, colleges, and universities in Texas; hence, students who permanently leave Texas attrit from the sample. I discuss the implications of this for identifica-

[19] Existing research finds mixed evidence on the degree to which parental job loss affects college enrollment. Hilger (2016) analyzes 7 million fathers' layoffs from 2000 to 2009 in the U.S. and finds paternal layoffs during adolescence caused only a 0.5 percentage-point reduction in college enrollment. On the other hand, Coelli (2011) studies a sample of Canadian adolescents from 1993 to 2007 and finds that parental layoffs were associated with to a 10 percentage-point decline in the probability of university enrollment.

In addition to the household income channel, parental job loss may also affect human capital investments by shifting student preferences. Ahlquist and Downey (2023) provide evidence parental exposure to the China shock caused children to be less likely to work in manufacturing, and Huttunen and Riukula (2019) find that paternal job loss caused children in Finland to be less likely to choose their father's field of study.

[20] Appendix Table 1.30 compares K-12 and college student demographics, K-12 spending, postsecondary tuition and appropriations, and high school and college graduation rates for Texas and the U.S. overall.

[21] Self-employed workers, independent contractors, many federal employees, members of the military, and participants in the informal sector are not covered by the state unemployment insurance system.

The records report employer industry codes (6-digit NAICS) in all years but do not report county of residency or county of employment.

tion in Section 1.4. Second, student linkages across time allow me to define exposure to the trade shock based on a student's county of residence prior to 2000 to avoid bias due to endogenous migration out of counties exposed to larger shocks. Thus, I limit my analysis samples to students observed prior to the start of the shock, which also bounds the time horizon over which treatment effects can be estimated. The panel begins with students in ninth grade in Fall 1995 and ends with those in ninth grade in Fall 2008.[22]

For "first-stage" estimations of the effects of PNTR on local labor market conditions and K-12 funding, I use data on employment and earnings from the Census County Business Patterns and Quarterly Workforce Indicators datasets, income data from the Bureau of Economic Analysis Regional Economic Accounts database, and school district finance data from the National Center for Education Statistics. I provide additional details on these supplemental data sources in Appendix 1.9.2.

1.4 Research Design

In this section, I describe my research design, identifying assumptions, and specification choices. I then provide empirical evidence supporting my identifying assumptions and characterize the "first-stage" effects of PNTR on local labor markets that may have affected students' educational choices.

1.4.1 Identification

PNTR's differential impact on local labor demand across counties motivates an identification strategy that compares the difference in outcomes of students in cohorts who made high school and college decisions after vs. before the policy change in counties that were more vs. less exposed to the labor demand shock. I follow Pierce and Schott (2020) in measuring county c's exposure to the establishment of Permanent Normal Trade Relations with China ($Exposure_c$, i.e., their treatment dose) as their employment-weighted average tariff gap using 1990 employment counts from the harmonized County Business Patterns Database (Eckert et al., 2021). I then discretize $Exposure_c$ into a binary treatment measure based on the population-weighted median tariff gap ($HighExposure_c$).[23] Figure 1.1 portrays variation in county-level exposure to the shock across the U.S. and Texas, and Table 1.1 presents descriptive statistics for samples of students from cohorts entering high school prior to 2000 across counties experiencing above- and below-median shocks.

Specifying treatment timing (τ) as when a cohort entered high school, I estimate the following event-study

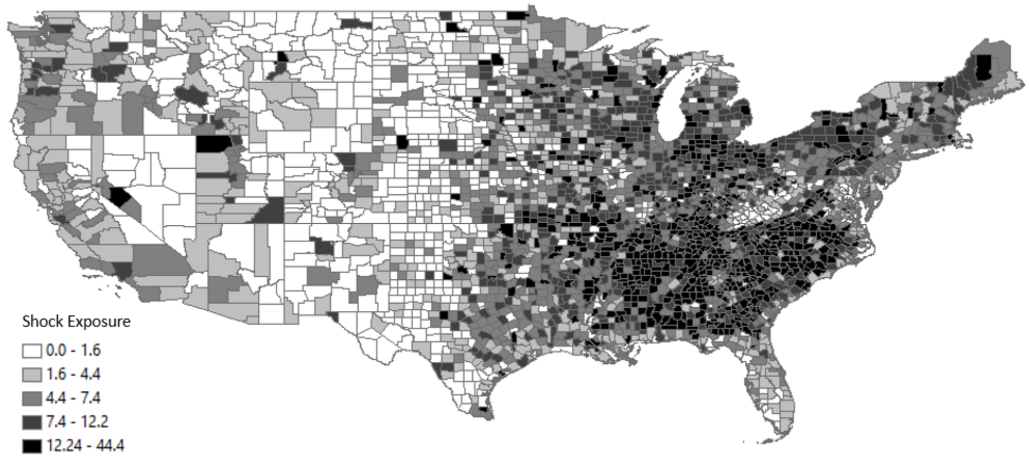
[22] My analysis sample ends with students in ninth grade in Fall 2008, because the youngest students that I can observe prior to the onset of the shock were kindergarteners in 1999, who would enter ninth grade in 2008. I further detail my sample selection in Appendix 1.9.2.

[23] Existing literature studying the local labor market effects of the China shock utilizes continuous treatment measures (e.g., Pierce and Schott, 2020), which impose a linear relationship between a county's employment-weighted tariff gap and student outcomes. However, the mechanisms through which exposure to import competition would affect educational decisions more plausibly conform with a threshold model. In other words, a *marginal* reduction in earnings for workers without a college degree may not induce additional students to enroll in college, but a *salient* decrease plausibly could.

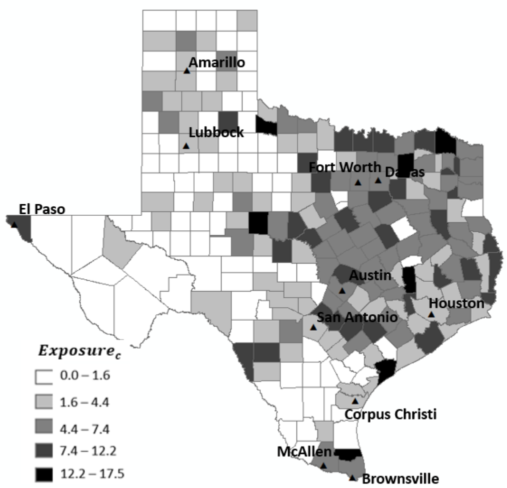
Table 1.1: Pre-Period Descriptive Statistics: Student Demographics & Outcomes

	(1) Below-Median Shock	(2) Above-Median Shock	(3) Diff
FRPL	0.470	0.416	0.054
	(0.499)	(0.493)	(0.046)
White	0.415	0.536	-0.121*
	(0.493)	(0.499)	(0.623)
Hispanic	0.424	0.289	0.135
	(0.494)	(0.453)	(0.079)
Black	0.132	0.149	-0.017
	(0.339)	(0.356)	(0.044)
Total AP/IB Courses	0.149	0.132	0.017
	(0.570)	(0.548)	(0.016)
Total Dual-Credit Courses	0.029	0.011	0.018
	(0.386)	(0.188)	(0.015)
Total Vocational Electvives	1.262	1.333	-0.071
	(1.451)	(1.461)	(0.090)
Total Industrial Electives	0.200	0.179	0.021
	(0.646)	(0.596)	(0.015)
Graduated HS	0.645	0.649	0.004
	(0.479)	(0.477)	(0.018)
Enrolled at Postsecondary Institution	0.392	0.387	0.005
	(0.488)	(0.487)	(0.014)
Enrolled at CTC	0.294	0.305	-0.011
	(0.456)	(0.460)	(0.013)
Enrolled at University	0.173	0.155	0.018*
	(0.378)	(0.362)	(0.010)
Certificate by 25	0.018	0.017	0.001
	(0.132)	(0.129)	(0.003)
Associate's by 25	0.037	0.039	-0.003
	(0.188)	(0.194)	(0.004)
Bachelor's by 25	0.121	0.117	-0.004
	(0.327)	(0.322)	(0.008)
Share of Qtrs Employed in TX at 30	0.519	0.505	0.014
	(0.471)	(0.472)	(0.016)
Unconditional Earnings at 30	22,602	21,375	1,227
	(34,371)	(35,913)	(921)
Employed All Qtrs in TX at 30	0.449	0.437	0.012
	(0.497)	(0.496)	(0.014)
Conditional Earnings at 30	47,064	45,661	1,403
	(37,655)	(37,843)	(1,831)
Observations	745,128	717,843	1,462,971

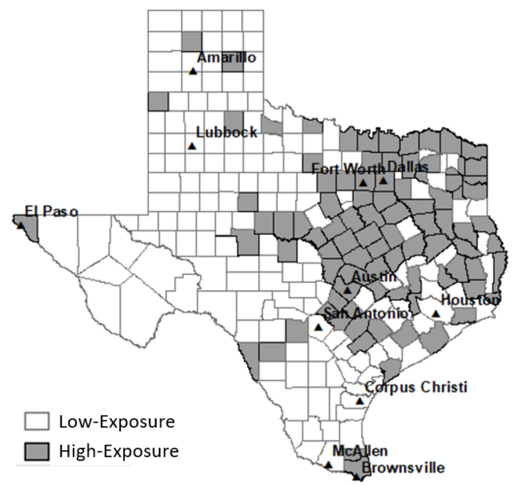
Notes: This table presents descriptive statistics for students from the 1995 to 1999 cohorts of ninth graders attending public schools in Texas. Students are divided by whether their county's $Exposure_c$ falls above or below the population-weighted median. College enrollment outcomes reflect enrollment in a public two- or four-year college or university in Texas within two years of expected high school graduation. Enrollment indicators for two- and four-year colleges are not mutually exclusive. Employment variables reflect employment in a position covered by unemployment insurance in Texas. College degree receipt is measured at age 25 and employment is measured at age 30.



(a) Shock Exposure, U.S.



(b) Shock Exposure, TX



(c) Binary Treatment Groups

Figure 1.1: County-Level Exposure to Local Shocks in the U.S. & Texas

This map county-level variation in exposure to adverse local shocks caused by Chinese import competition, as measured by a county's employment-share weighted "tariff gap" across the industries (i.e., product specialties) present in the county. Each industry's tariff gap is defined as the difference between preferred tariff rates locked in by the establishment of Permanent Normal Trade Relations with China in 2000 and punitive import tariff rates set by the Smoot-Hawley Tariff Act of 1930. Following Pierce and Schott (2020), I define employment shares based on 1990 Census County Business Patterns data and industry-level tariff gaps are measured in 1999. Panel (a) divides counties into quintiles of exposure for the entire U.S. and panel (b) zooms in to Texas. Panel (c) groups counties into binary high-exposure and low-exposure groups based on the population-weighted median.

specification:[24]

[24] I do not specify treatment centered around an earlier grade because the Texas Child Labor Law first allows children to work (subject to hour restrictions) at 14, the age at which students commonly enter ninth grade. Thus, there are no market-defined opportunity costs of schooling before this age. On the other hand, I do not specify treatment centered around a later grade, because I am interested in adjustment mechanisms that occur during high school: vocational, AP, and dual-credit course-taking, along with diploma receipt.

$$y_{ict} = \sum_{\substack{\tau=1995 \\ \tau \neq 1999}}^{2008} \pi_{\tau}^{ES} \mathbf{1}\{t = \tau\} * HighExposure_c + \alpha_c + \alpha_t + \Phi S_i + \sum_{\substack{\tau=1995 \\ \tau \neq 1999}}^{2008} \Gamma_{\tau} \mathbf{1}\{t = \tau\} \times X_c + \Theta Z_{ct} + \varepsilon_{ict} \quad (1.1)$$

In equation (1.1), I regress outcome y_{ict}, such as an indicator variable for graduating high school, for student i in county c and ninth-grade cohort t on $HighExposure_c$ interacted with an indicator variable for belonging to a particular cohort τ, along with county (α_c) and cohort (α_t) fixed effects that absorb county-specific characteristics common across all cohorts and cohort-specific characteristics common across all counties, respectively. In addition to time-invariant student demographics (S_i), I control for two sets of county-level covariates.[25] First, I include a vector of pre-period measures of county economic and demographic characteristics (X_c) interacted with cohort dummies to account for time-varying shocks related to a county's economic profile; second, I control for a vector of other trade policy changes that may affect local labor demand (Z_{ct}).[26] Because X_c includes a county's baseline manufacturing share, my identifying variation consists of comparisons of students in counties with similar *overall* manufacturing presence but with differing shares of employment in firms specializing in *specific* products that were exposed to import competition. π_{τ}^{ES} (Event Study) represents the difference in outcomes between such students among cohort τ relative to this difference among the 1999 ninth grade cohort.

Although $HighExposure_c$ is a binary measure, the underlying exposure to import competition is inherently continuous. Equation (1.1) thus compares changes in outcomes across cohorts of students from counties exposed to exogenously "big" and "small" labor demand shocks without a group of "pure" control counties that are completely unaffected by PNTR. Such comparisons cannot identify the causal effect of a labor demand shock relative to the shock's absence (Callaway et al., 2021). However, they can identify the causal effect of exposure to a large local labor demand shock relative to a small one. Identification requires the assumption that trends in outcomes of students high-exposure counties would on average evolve in parallel to those of students from low-exposure counties if all counties had instead experienced a small shock.[27] In

[25] S_i consists of indicator variables for student race and ethnicity, gender, Limited English Proficiency status, and free-or-reduced-price lunch eligibility. These characteristics are included as controls to reduce residual variation, but I show that results are robust to their exclusion in Section 1.7.

[26] X_c consists of 1990 measures of median household income, population share without a college degree, the foreign-born population share, and the share of employment in manufacturing, along with the per capita volume of shale oil and gas reserves within a county's borders. Z_{ct} includes the average import tariff rate associated with a county's goods, the county's exposure to the end of global restrictions on textiles and clothing imports from the phasing out of the Multi-Fiber Arrangements, and the county's exposure to changes in tariffs on imports into China and Chinese domestic production subsidies. Both sets of covariates are adopted from Pierce and Schott (2020), with the addition of shale oil and gas reserves to account for the fracking boom – a potential confounder specific to my setting of Texas that Kovalenko (2023) shows caused students to leave school early to enter newly robust local labor markets. I show in Section 1.7 that my results are robust to excluding both sets of controls from the specification.

[27] Specifically, under this assumption, such comparisons identify the average causal effect of exposure to a large local labor demand shock relative to a small one for population of students from high-exposure counties – analogous to an Average Treatment Effect on the Treated parameter. Under the additional assumption that on average a high treatment dosage would have the same effects on students from both groups, equation (1.1) identifies the average causal effect of exposure to a large local labor demand shock relative to a small one for all Texas students – analogous to an Average Treatment Effect parameter (Callaway et al., 2021).

other words, one must assume that a low treatment "dosage" (i.e., a small decrease in labor demand) would have the same effects on students from both treatment groups.

Estimates of equation (1.1) with outcomes proxying for labor demand on the left-hand side (e.g., per-capita labor income) inform the appropriate estimator for identifying the causal effects of exposure to negative local labor demand shocks on human capital accumulation and later-life earnings. Event studies in Section 1.4.2 show that employment and earnings grew faster leading up to the policy change in counties that were more exposed to PNTR-induced labor demand shocks than in less-exposed counties. These patterns suggest that student outcomes would not have trended in parallel across more- and less-exposed counties if both groups were exposed to a low treatment dosage, but more likely would have continued along existing differential trends.[28] Therefore, I explicitly control for differences in existing trends in outcomes using the following two-step procedure proposed by Goodman-Bacon (2021c):[29]

$$y_{ict} = \lambda t * HighExposure_c + \alpha_c + \alpha_t + \Phi S_i + \sum_{\tau=1995}^{1998} \Gamma_\tau \mathbf{1}\{t=\tau\} \times X_c + \Theta Z_{ct} + \varepsilon_{ict} \tag{1.2}$$

$$\widetilde{y}_{ict} = \sum_{\substack{\tau=1995 \\ \tau\neq 1999}}^{2008} \pi_\tau^{DTES} \mathbf{1}\{t=\tau\} * HighExposure_c + \alpha_c + \alpha_t + \Phi S_i + \sum_{\substack{\tau=1995 \\ \tau\neq 1999}}^{2008} \Gamma_\tau \mathbf{1}\{t=\tau\} \times X_c + \Theta Z_{ct} + \varepsilon_{ict} \tag{1.3}$$

In the first step (equation (1.2)), I estimate a linear pre-trend in outcomes by regressing the outcome of interest y_{ict} on a linear trend interacted with $HighExposure_c$, using only cohorts entering ninth grade prior to 2000. I then extrapolate the estimated differential pre-trend ($\hat{\lambda}$) beyond 2000 and construct $\widetilde{y}_{ict}$, the *de-trended* outcome variable, by partialing out $\hat{\lambda} t * HighExposure_c$.

In the second step (equation (1.3)), I estimate the same event study specification as equation (1.1) but with the de-trended outcome variable ($\widetilde{y}_{ict}$) on the left-hand side. π_τ^{DTES} (De-Trended Event Study) are the differences in outcomes between students from more-exposed and less-exposed counties among cohort τ relative to the extrapolated linear pre-trend.[30] Following Kuka et al. (2020), I construct standard errors using

[28] A particular threat to the modified parallel trends assumption is that continued differential growth of labor market opportunities in high-exposure counties relative to low-exposure counties could increase the opportunity costs of schooling, such that educational attainment would fall. On the other hand, this differential growth could increase educational attainment by increasing parental income. In either case, the standard difference-in-differences equivalent of equation 1.1 would not identify causal effects.

[29] This two-step approach and the similarly spirited "parametric event study" specification proposed by Dobkin et al. (2018) have been used to estimate causal effects of hospital admissions on financial health (Dobkin et al., 2018), bankruptcy flag removal on consumer spending (Gross et al., 2020), bankruptcy reform on credit behavior, school finance reforms on student achievement (Lafortune et al., 2018), abortion denial on financial health (Miller et al., 2023), opioid supply on local economic conditions (Beheshti, 2022), immigration reform on educational attainment (Kuka et al., 2020), and Paycheck Protection Program loans on employee retention (Autor et al., 2013).

[30] An event-study plot of π_τ^{DTES} across cohorts intuitively resembles that of π_τ^{ES} with its pre-trend rotated toward the x-axis.

a degrees-of-freedom adjustment to account for utilizing a regression-adjusted outcome variable.

The two-step event study estimates are useful for showing visually whether a linear pre-trend fits the data and assessing dynamic responses to treatment. To summarize treatment effects across cohorts, I estimate a difference-in-differences specification as the second step, replacing equation (1.3) with the following equation:

$$\widetilde{y}_{ict} = \Pi^{DTDD} HighExposure_c * \mathbf{1}\{t \geq 2000\} + \alpha_c + \alpha_t + \Phi S_i + \Gamma_\tau \mathbf{1}\{t \geq 2000\} \times X_c + \Theta Z_{ct} + \varepsilon_{ict} \qquad (1.4)$$

where I interact $HighExposure_c$ (and control measures, X_c) with an indicator variable for entering high school after the start of the shock instead of with individual cohort dummy variables. Identification of the causal effect of exposure to the local labor demand shock (Π^{DTDD}; De-Trended Difference-in-Differences) requires the assumption that differences in outcomes of students from more- and less-exposed counties would continue to evolve along the existing linear trend if they both were exposed to small labor demand shocks.

Greenland et al. (2019) and Burga and Turner (2022) find evidence of the China shock increasing out-migration from exposed local labor markets, particularly among young adults and the less-educated. Such out-migration among lower-ability students would yield selection into treatment in a naive framework based on contemporaneous residence, biasing estimates of the effects of exposure to the labor demand shock on educational attainment and later-life earnings upward. Therefore, I assign treatment status and county-level covariates to students based on their county of residence *prior* to the onset of the shock in an Intent-to-Treat (ITT) framework. This approach is made possible by my ability to link students across years and datasets in the individual-level UHERC data. Because I cannot observe students that leave Texas, I also test for differential attrition rates (Figure 1.8) and find no statistically significant evidence of differential attrition by treatment status at any age from 16 through 30.[31]

I note two conditions that are sufficient for the required assumption that differences in student outcomes across more- and less-exposed counties would continue along existing linear trends if both groups experienced small labor demand shocks to hold. First, one must assume that such differences in student outcomes would continue along existing trends in the absence of the policy change (Goodman-Bacon, 2021c). Con-

[31]Figure 1.8 presents separate estimates of equation (1.4) with indicator variables equal to one if an individual is not observed in the data again from that age through age 30 as the outcomes. The 95% confidence intervals for every estimate include 0. If the negative point estimates are taken at face value, they suggest exposure to the labor demand shock caused students to be *less* likely to leave Texas. For such behavior to bias estimates of the effect of exposure to the shock on human capital accumulation and later-life earnings upward, these marginal "stayers" would need to be positively selected. A plausible story conforming with this notion would be that students that otherwise would have attended out-of-state colleges and found work outside of Texas instead attended in-state schools because of declines in family income caused by the shock. I provide evidence against this story in Section 1.7.

The increase in magnitude of estimates in 1.8 as the age of definition approaches 30 is largely a mechanical effect. Because I observe no workers past age 30, "attrition" at age 29 only reflects non-participation in the labor force or education system in Texas for 2 years, as opposed to non-participation for 10 years for attrition defined at age 20. To the extent that human capital adjustments increased labor force opportunities, increases in employment for a given year appear as reduced attrition for later ages, while more weakly relating to earlier defined measures.

sistent with this condition, I show that differences in *pre-measured* characteristics of ninth-grade students do not deviate from existing trends following the policy change. Table 1.2 presents estimated "effects" of exposure to larger shocks on baseline (pre-shock) characteristics using equation (1.4) and shows no evidence of local shocks "affecting" previously measured student race, ethnicity, gender, English-Language Learner status, free-or-reduced-price-lunch eligibility, or an index of later-life earnings predicted on these measures. Exposure to larger shocks correlates with a statistically insignificant $62 (2020$) decrease in earnings predicted on fixed student characteristics, suggesting that earnings would have continued along existing trends in the absence of the labor demand shock. Second, one must assume a small shock to local labor demand would cause the same average deviations from existing differential trends for students from high-exposure and low-exposure counties (Callaway et al., 2021). Table 1.1 presents average characteristics of individuals from high-exposure and low-exposure counties and shows the groups only statistically differ from one another along 2 out of 19 dimensions. The balance in observable characteristics of indiviudals across both sets of counties suggests that members of each treatment group would respond similarly to the same sized labor demand shock.

Table 1.2: Specification Tests: Local Shocks Did Not "Affect" Pre-determined Characteristics

	(1) White	(2) Hispanic	(3) Black	(4) Male	(5) ELL	(6) FRPL Eligiblity	(7) Predicted Earnings
De-trended DD	0.003	-0.000	-0.006	0.003	0.008	0.022	-62
	(0.007)	(0.007)	(0.008)	(0.005)	(0.011)	(0.023)	(76)
Percent Change	0.6	-0.1	-4.0	0.6	7.3	5.1	-0.1
Pre-period Mean	0.516	0.325	0.137	0.514	0.109	0.435	48,265
N	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707

DoF-adjusted standard errors in parentheses are clustered at the county level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates from falsification exercises of the "effects" of exposure to local shocks on pre-determined student demographics, using individual-level data from the University of Houston Education Research Data. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c) The specification controls for county and year fixed effects, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset and assigned demographic measures from their first appearence in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. The outcome variables across columns are as follows: (1) an indicator for reporting white as race and non-Hispanic as ethnicity, (2) an indicator for reporting Black as race and non-Hispanic as ethnicity, (3) an indicator for reporting Hispanic as ethnicity, (4) an indicator for reporting male as sex, (5) an indicator for classification as an English-Language Learner, and (6) an indicator for Free-or-Reduced-Price-Lunch eligibility. The outcome in column (7) is a predicted index of earnings at age 30 constructed by estimating the relationship between earnings and student demographics with a sample of pre-period cohorts and then predicting later-life earnings for all cohorts based on these relationships. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

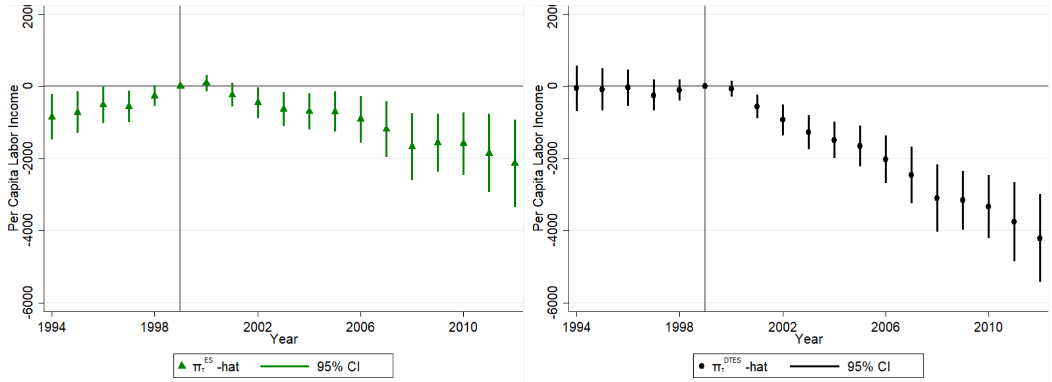
In Section 1.7, I show that my main results are generally robust to two relaxations of the above assumptions. First, I allow potential outcomes of students from high-exposure and low-exposure counties to revert to trending in parallel after continuing along existing differential linear trends for only p periods, rather than the entire sample period. In practice, this exercise makes comparisons of differences in outcomes from high-exposure and low-exposure counties relative to a pre-trend that flattens off after p periods, instead of relative to the continuation of $\hat{\lambda}$ throughout the post-period. Second, I allow for smooth deviations in parallel outcomes from the existing linear pre-trends, rather than assuming exact linearity (Rambachan and Roth, 2023).

1.4.2 Effects of PNTR on Local Labor Markets

Using data from the County Business Patterns Database (Eckert et al., 2021), Bureau of Economic Analysis Regional Economic Accounts, and Census Quarterly Workforce Indicators, I characterize the local labor demand shock caused by PNTR in Texas before examining student responses. This serves dual purposes of assessing the relevance of the specified exposure measure and guiding interpretation of the channels through which the China shock may have affected student behavior. In an initial example, I present results from three specifications that build toward causal identification: standard event-study estimates (equation 1.1), de-trended event-study estimates (equation 1.3), and de-trended difference-in-difference estimates (equation 1.4) – my primary measure of treatment effects. Afterward, I only present de-trended event studies and difference-in-difference estimates in the main text and present standard event studies in Appendix 1.9.1.1.

I first examine how exposure to PNTR affected local income, a proxy for overall labor demand. Panel (a) of Figure 1.2 presents estimates of π_τ^{ES} that represent the difference in per-capita wage and salary income in high-exposure versus low-exposure counties relative to this difference in 1999. Income grew faster in counties that were later more exposed to PNTR than in less-exposed counties leading up to the policy change. The continuance of this dynamic if both groups were exposed to small labor demand shocks would violate the modified parallel trends assumption and bias the estimated effect of greater exposure to import competition on local income upward. Therefore, I explicitly control for the difference in existing trends using equation 1.3. Panel (b) presents estimates of π_τ^{DTES}, representing the difference in per-capita labor income between high-exposure and low-exposure counties in each year relative to the evolution implied by the differential growth before 2000. Coefficients for before 2000 hover around 0, confirming the fit of a linear trend and supporting the validity of the specification. Under the assumption that earnings in more-exposed counties would continue to grow relative to less-exposed counties along this existing trend, post-period coefficients suggest that greater exposure to import competition dramatically reduced income. The difference-in-difference estimate (π^{DTDD}) in Table 3.8 summarizes this effect, indicating that PNTR caused a statistically significant 18% decline ($p < 0.01$) in earnings in high-exposure counties.

In addition, Table 1.3 and Figure 1.3 provide evidence of the labor demand shock's effect on employment. Difference-in-differences estimates indicate that greater exposure to import competition caused the loss of an average of approximately 1,400 manufacturing jobs per county and reduced employment relative to popula-



(a) Per-Capita Labor Income

(b) Adjusted for Linear Pre-trends

Figure 1.2: Import Competition Caused Declines in Local Labor Demand

Notes: These figures present estimates of the effect of exposure to Chinese import competition on per-capita labor income using personal income data from the Bureau of Economic Analysis Regional Economic Accounts dataset and population counts by age group from the Survey of Epedemiology and End Results. Estimates in panel (a) reflect coefficients from event study regressions (equation (1.1)) that compare the difference in outcomes between counties with above- and below-median exposure to local labor market shocks caused by import competition over time relative to this relationship in 1999, one year prior to the start of treatment. Estimates in panel (b) reflect coefficients from a the two-step de-trended event study (equation (1.3)) that partials out a linear pre-trend in the first step. Both specifications control for county and year fixed effects, 1990 county characteristics flexibly interacted with year dummies, county exposure to the fracking boom, and other changes to US trade policy. Standard errors are clustered by county. Standard errors for the two-step procedure reflected in panel (b) account for parameters estimated in the first step via a degrees-of-freedom adjustment.

Table 1.3: Effects of Import Competition on Local Employment & Income in Texas

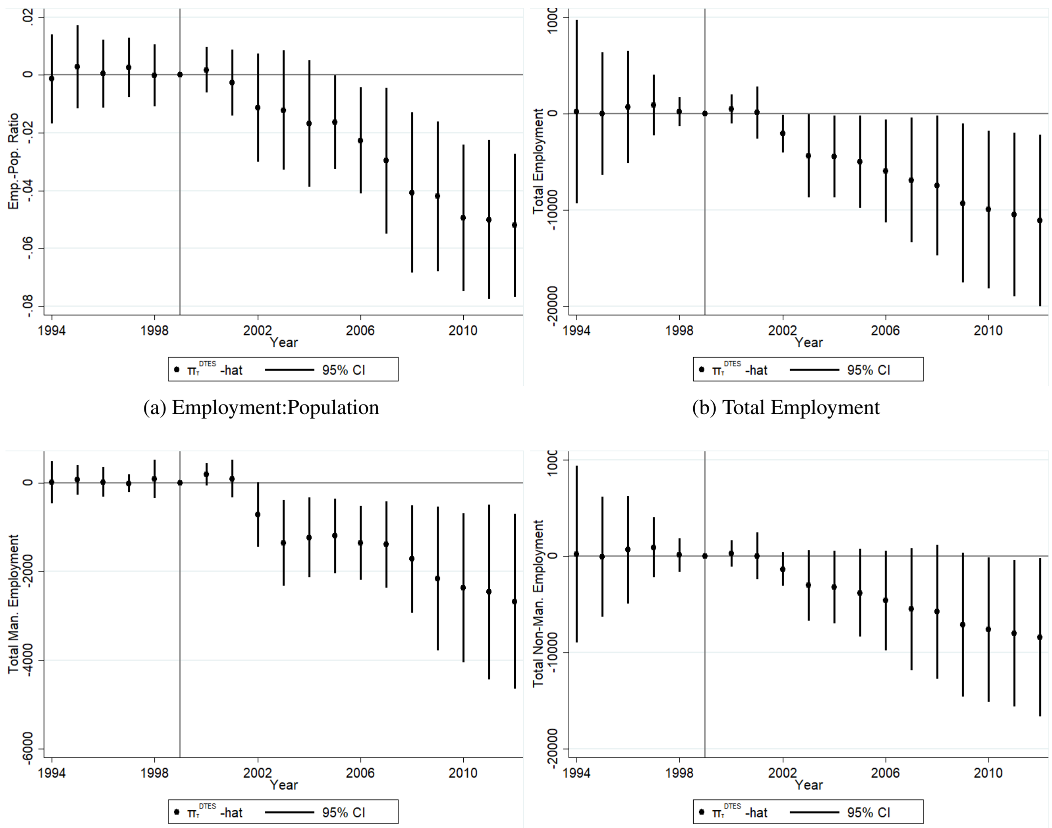
	(1) Earnings Per Capita	(2) Man. Emp	(3) Total Emp	(4) Emp:Pop
De-trended DD	-2,194***	-1,443***	-5,370*	-0.030***
	(258)	(462)	(3,158)	(0.008)
Percent Change	-17.6	-22.3	-13.2	-6.8
Pre-period Mean	12,432	6,475	40,702	0.436
N	4,810	4,810	4,810	4,810

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effects of exposure to Chinese import competition on employment-to-population ratios and per capita income, using employment counts from the County Business Patterns Database, population counts by age group from the Survey of Epedemiology and End Results, personal income data from the Bureau of Economic Analysis Regional Economic Accounts dataset, and employment and wage data broken down by age and education levels from the Quarterly Workforce Indicators dataset. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between county-level outcomes in more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Standard errors are clustered by county and adjusted to account for using a regression-adjusted estimator. The outcome variables across columns are as follows: (1) per-capita wage and salary income, manufacturing (2) and overall employment (3) in the county, and (4) employment-to-population ratio. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

tion by 3.0 percentage-points. Both coefficients are statistically significant ($p < 0.01$), and the magnitudes are similar to previous estimates of the effects of the China shock on local labor markets using nationwide samples and nearly identical to the effects of a recession on more- exposed local labor markets.[32] Moreover, the point estimate in column (3) suggests that the negative employment effects of import competition spilled over outside of manufacturing.[33]



(a) Employment:Population

(b) Total Employment

(c) Manufacturing Employment

(d) Non-Manufacturing Employment

Figure 1.3: Import Competition Caused Declines in Manufacturing and Overall Employment

Notes: These figures present estimates of the effect of exposure to Chinese import competition on employment using employment counts from the County Business Patterns Database, population counts by age group from the Survey of Epedemiology and End Results, and personal income data from the Bureau of Economic Analysis Regional Economic Accounts dataset. Estimates reflect coefficients from two-step event study regressions that compare the difference in outcomes between students from counties that were more or less exposed to negative labor demand shocks across cohorts relative to the continuation of a linear pre-trend (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, 1990 county characteristics flexibly interacted with year dummies, county exposure to the fracking boom, and other changes to US trade policy. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable.

[32] Autor et al. (2021) finds a 1.9 percentage-point decline in employment relative to population in exposed local labor markets. Hershbein and Stuart (2023) show that each recession since the 1970s corresponded to approximately a 3 percentage-point employment decline in counties with above-median exposure. Although Autor et al. (2021) shows that the effects of Chinese import competition persisted for nearly 20 years after the onset of the shock, much of the China shock literature examines effects on local labor market outcomes only through 2007. In Appendix Table 1.13, I show robustness to setting the end of my panel at 2007, 2012, and 2016.

[33] Existing literature on the China shock finds mixed evidence on how the shock affected employment outside of manufacturing. Using a shift-share design and aggregate Census data, Autor et al. (2013) find that local labor markets (defined as commuting zones) exposed to the China shock experienced decreases in non-manufacturing employment relative to population, and this reduction statistically differs from zero for workers without a college education. Bloom et al. (2019) utilize the same design and establishment-level data from the Census Bureau's Longitudinal Business Database and find commuting-zone-level exposure to the China shock resulted in small increases in non-manufacturing employment. Both papers find reductions in manufacturing and overall employment in exposed local labor markets.

Declines in local labor demand may enter into educational decisions by decreasing the opportunity cost of schooling. I define two proxies for opportunity costs: the average earnings for school-aged (15-24) workers and the average earnings for workers who never attended college. Table 1.4 shows that the labor demand shock significantly reduced both measures by 18% and 8%, respectively. To the extent that students expected the declines in labor market opportunities for workers without college experience to persist, the second estimate also represents a decrease in the expected lifetime earnings associated with entering the labor market before or directly after high school completion. The estimate in Column (3) indicates that the decrease in earnings for workers without college experience translated to an increase in the college earnings premia. Both the decline in opportunity costs and the increase in the labor market return to a college education would incentivize marginal students to enroll in college.

Table 1.4: Relevance of the Labor Demand Shock to Human Capital Decisions

	(1) Earnings, 14-24	(2) Earnings, No College	(3) College Earnings Premia	(4) Per-pupil Property Tax Revenue	(5) Per-pupil K-12 Spending	(6) FRPL Eligibility
De-trended DD	-3,620***	-2,839***	0.257***	-2,786***	39	0.030***
	(365)	(477)	(0.025)	(421)	(98)	(0.010)
Percent Change	-18.0	-8.2	15.0	-75.7	0.43	8.0
Pre-period Mean	20,166	34,621	1.709	3,681	9,094	0.370
N	4,286	4,298	4,302	4,810	4,810	3,678,707

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effects of exposure to Chinese import competition on earnings by age and educational attainment, K-12 spending, and eligibility for Free-or-Reduced-Price Lunch using wage data broken down by age and education levels from the Quarterly Workforce Indicators dataset, K-12 finance data from the National Center for Education Statistics Common Core of Data, and student-level data from the UHERC. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between county-level (and student-level) outcomes in more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Column (6) further controls for student demographics, including pre-shock FRPL-eligibility. Standard errors are clustered by county and adjusted to account for using a regression-adjusted estimator. The outcome variables in columns (1) and (2) are average earnings for school-aged workers and workers without any college attainment, respectively. The outcome in column (3) is the ratio of earnings for workers with a college degree to those without any college attainment. The outcomes in columns (4) and (5) are per-pupil school district revenue from local property taxes and per-pupil current expenditure, respectively. The outcome in column (6) is an indicator variable for eligibility for the Free-or-Reduced-Price Lunch program. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

On the other hand, local economic shocks also may affect human capital accumulation by decreasing family income and school funding (Stuart, 2022; Burga and Turner, 2022). If property values fell in exposed counties, as shown with a nationwide sample by Feler and Senses (2017), accompanying reductions in school spending could negatively affect short- and long-run student outcomes.[34] However, Texas' K-12 finance system provides insurance to local fluctuations in property tax revenues through its "Robin Hood" formula, which redistributes excess property tax revenues to equalize per-pupil spending across districts. Columns (4)

[34]For a review of recent literature on the causal effects of K-12 school spending, see Jackson (2018).

and (5) of Table 1.4 show that despite district revenues from local property taxes falling by approximately \$2,800 per-pupil in more-exposed counties, school spending was unaffected.[35] Appendix Table 1.31 supports that the "insurance" provided by the Robin Hood formula is unique relative to the rest of the country: while school spending in Texas fell by 0 cents per dollar lost in district property tax revenues due to local shocks, in the rest of the country, school spending fell by 84 cents per dollar lost in local revenues. Still, exposed students on average experienced declines in family income, as evident by an increase in eligibility for free-or-reduced-price lunch shown in Column (6).

All together, these results support that, in Texas, marginal students experienced increased incentives to acquire more education. Although they did not see reductions in the "external" public investments that support such attainment (i.e., K-12 funding), they may have experienced declines in "external" private investments (i.e., family expenditures). Thus, the direction of the effect of the labor demand shock on overall human capital accumulation is theoretically ambiguous.

1.5 Main Results

My primary analyses examine how students adjusted their educational decisions in response to the labor demand shock and whether these adjustments shielded students from the shock's negative effects on earnings.

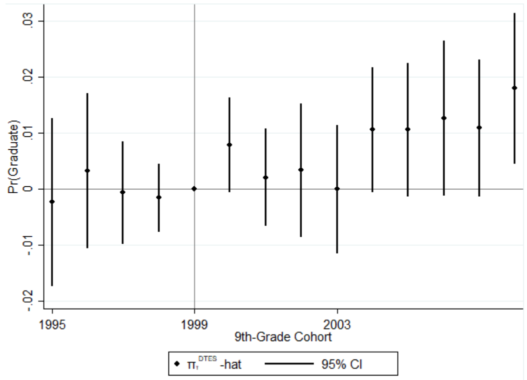
1.5.1 Human Capital Adjustments in High School

The first-stage results discussed in Section 1.4.2 are consistent with the China shock incentivizing students on the margin of dropping out of high school to instead graduate by reducing the opportunity cost of schooling and increasing the returns to higher levels of education. However, existing estimates of the effects of the China shock on high school graduation rates with nationwide samples give mixed evidence on whether such responses occurred. (Greenland and Lopresti, 2016; Burga and Turner, 2022).[36] Figure 1.4a and Table 1.5 present de-trended event study and difference-in-differences estimates of the effects of local shock exposure on high school graduation. I find no statistically distinguishable effects across all students and can rule out effects larger than a 2.0 percentage-point increase in the probability of high school graduation. The null result is consistent with the findings of Burga and Turner (2022), and my intent-to-treat specification is not subject to the potential bias from student migration that they caution threatens identification of the China shock's effects on high school graduation.[37]

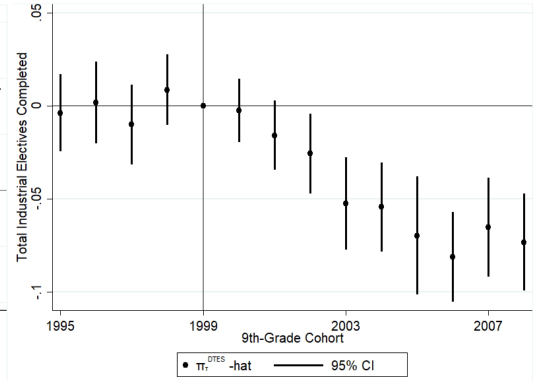
[35] Appendix Figure 1.11 presents the corresponding event studies.

[36] Greenland and Lopresti (2016) and Burga and Turner (2022) both examine the effects of the China shock on high school graduation rates using aggregated graduation counts for nationwide samples but come to different conclusions. Greenland and Lopresti (2016) find that graduation rates increase in local labor markets exposed to the China shock by 3.6 percentage-points; however, Burga and Turner (2022) provide evidence that this result is mostly explained by outmigration and weak instrument bias.

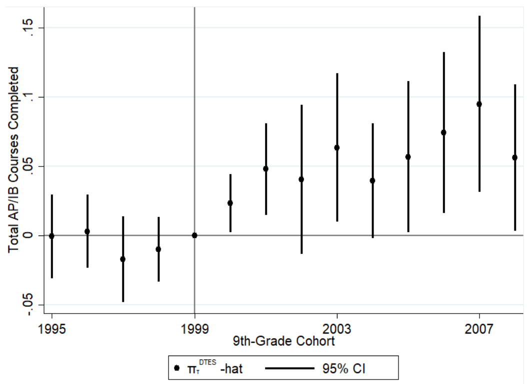
[37] Identification with my ITT specification still would be threatened by exposure to the China shock increasing the likelihood that families migrated out of Texas, altogether. However, in Section 1.7, I test for differential attrition from Texas and find no statistically significant evidence of such behavior.



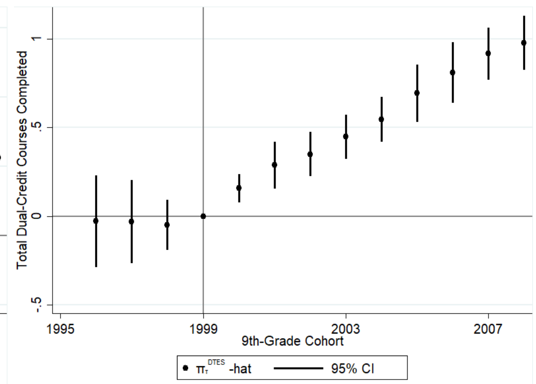
(a) HS Graduation



(b) # of Industrial Electives



(c) # of AP/IB Courses



(d) Dual Enrollment

Figure 1.4: Effects of the Labor Demand Shock on Human Capital Accumulation in HS

Notes: These figures present estimates of the effect of exposure to negative labor demand shocks during youth and adolescence on high school graduation and course selection. Estimates reflect coefficients from two-step event study regressions that compare the difference in outcomes between students from counties that were more or less exposed to negative labor demand shocks across cohorts relative to the continuation of a linear pre-trend (Goodman-Bacon, 2021c). The outcomes are (a) an indicator for graduating high school, counts of the number of (b) manufacturing-aligned electives and (c) AP/IB courses completed, and (d) an indicator variable for enrolling in a dual-credit course at a local college. The specification controls for county and year fixed effects, student demographics, 1990 county characteristics flexibly interacted with year dummies, county exposure to the fracking boom, and other changes to US trade policy. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable.

Null effects on high school graduation do not rule out important *intensive-margin* adjustments students could make in response to the shock. Vocational courses can give students specialized human capital that prepares them to find employment in a particular industry, and research links participation in such courses with higher earnings (Bishop and Mane, 2004). However, the loss of manufacturing jobs in counties exposed to import competition likely reduced the labor market return to completing vocational courses in manufacturing, in particular. Thus, I examine whether exposure to the labor demand shock affected overall and

Notably, the coefficients in Figure 1.4a representing the effects of exposure to the labor demand shock on high school graduation rates among cohorts entering high school between 2004 and 2008 are all marginally significant and fall between 1.0 and 1.9 percentage point increases. The 95% confidence interval for each coefficient still rule out the 3.6 percentage-point increase found by Greenland and Lopresti (2016), potentially due to the use of different estimators, differences in responses in Texas and the rest of the country, or selective outmigration reflected in their aggregated graduation counts.

Table 1.5: Effects of the Labor Demand Shock on HS Graduation and Course Selection

	(1) HS Grad	(2) Total Voc. Electives	(3) Industrial Electives	(4) Dual-Enrollment Courses	(5) AP/IB Courses
De-trended DD	0.008	-0.039	-0.047***	0.581***	0.059**
	(0.006)	(0.038)	(0.012)	(0.095)	(0.023)
Percent Change	1.2%	-2.8	-24.3	299.3	40.3
Pre-period Mean	0.706	1.398	0.191	0.194	0.147
N	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707

DoF-adjusted standard errors in parentheses are clustered at the county level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effects of exposure to negative local labor demand shocks on high school educational attainment and course selection using individual-level data from the University of Houston Education Research Data. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. The outcome in column (1) is an indicator for graduating high school. The outcomes in columns (2) - (4) are counts of the total vocational elective courses, industrial elective courses, and dual-credit courses completed in high school. The outcome column (5) is the total number of AP or IB courses completed in high school. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for treated counties.

manufacturing-specific vocational course-taking in high school using Texas Education Agency field classifications.[38] Table 1.5 shows no evidence of changes in overall vocational course-taking, but that consistent with salient reductions in local demand for manufacturing workers, students enrolled in 24% fewer industrial electives ($p < 0.01$). Appendix Table 1.16 presents estimates of the effects of shock exposure on course-taking across other individual categories of vocational electives. Students exposed to larger shocks took significantly fewer technology-based electives and more agricultural and business electives.

Students may also actively prepare for pursuing a postsecondary education while in high school by taking courses eligible for college credit. I test if students responded to salient increases in the college earnings premium by doing so and present results in in Figure 1.4 and columns (4) and (5) of Table 1.5. I find that students exposed to larger local shocks completed 40% more Advanced Placement or International Baccalaureate courses and 0.5 (299%) more courses through dual-enrollment with local colleges.[39] These estimates suggest the reduction in labor market opportunities for workers without college experience increased desire to pursue a college education. Moreover, Jackson (2010) provides evidence that AP course-taking increases the likelihood of college matriculation, such that the estimated increases in AP course-taking and dual enrollment also represent mechanisms that may have aided students in reaching college as a manner of adjustment

[38] Appendix 1.9.2.3 details each elective category.

[39] The magnitude of the estimated effect on dual-credit completions should be interpreted with caution when compared to the pre-period mean, because overall take-up of dual-credit increased four-fold during the sample period.

to the labor demand shock.

1.5.2 Postsecondary Human Capital Adjustments

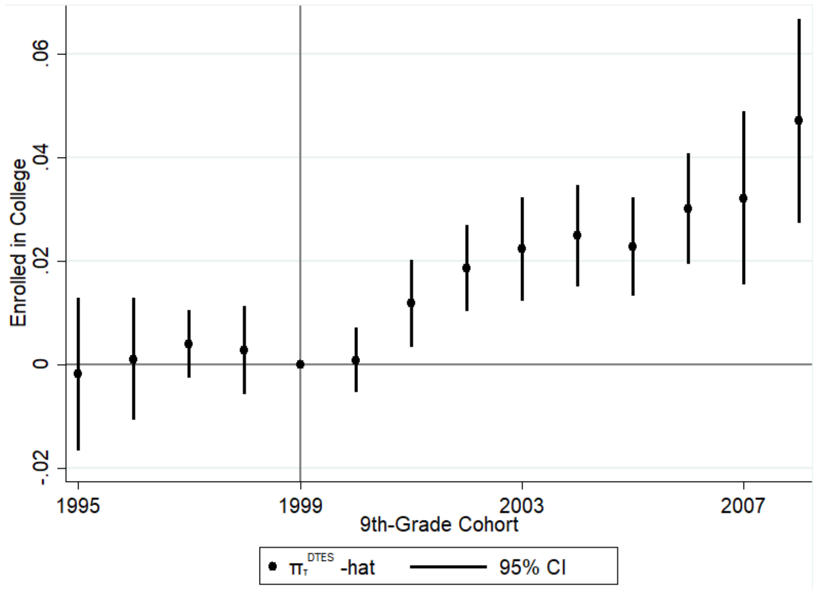
College plausibly offered a critical avenue for students to shield themselves from the labor demand shock. First-stage estimates from Section 1.4.2 suggest that earnings for workers with a college degree increased relative to those without college experience as a result of the shock, and existing work provides evidence of considerable returns to both two-year and four-year degrees even for marginal students (e.g., Card, 2001; Smith et al., 2020). However, the substantial monetary costs of higher education in combination with the shock's negative effects on household income may have prevented the above increases in AP and IB course-taking and dual enrollment from translating into college matriculation by marginal students.

Figure 1.5 presents de-trended event study estimates of college enrollment within two years of expected high school completion that suggest students were able to adjust to the shock by pursuing a higher education at two- and four-year institutions. De-trended difference-in-differences estimates in columns (1) through (3) of Table 1.6 indicate that students exposed to local shocks were 1.8 percentage-points (4) more likely to enroll in any Texas public college, 1.2 percentage-points (4%) more likely to enroll in a two-year college, and 1.6 percentage-points (10%) more likely to enroll at a four-year university.[40] Each estimate is statistically significant ($p < 0.01$), and the magnitudes are comparable to existing estimates of the effects of placement in smaller elementary school classrooms (Chetty et al., 2011) or receipt of $1,000 in additional financial aid (Castleman and Long, 2016).[41] Because I only observe whether students enroll in in-state colleges and universities, a notable concern is that the estimated increases in enrollment may in part reflect substitution from more expensive out-of-state to cheaper in-state institutions in response to the shock's negative effect on family income. However, in Section 1.7, I find no evidence of substitution in enrollment from private to public in-state universities or from higher- to lower-priced in-state institutions.

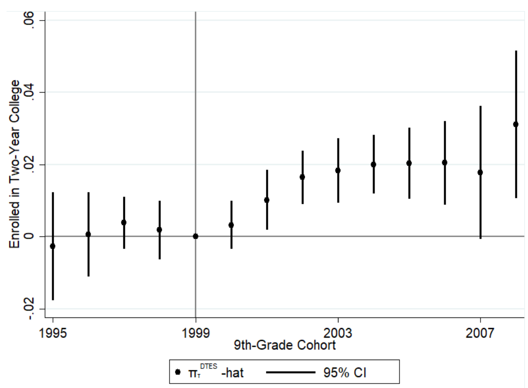
The above increase in college enrollment may not have improved labor market outcomes for students if they did not persist toward degree receipt. Thus, I use two exercises to assess whether exposure to local shocks caused meaningful increases in postsecondary educational attainment beyond initial college enrollment. First, I separately estimate the effects of exposure to local shocks on college enrollment – including dual enrollment while in high school – at each age from 16 through 30. Figure 1.6 presents coefficients from de-trended difference-in-difference specifications estimated separately for each age. Each coefficient from

[40] Because I define enrollment outcomes based on attending a college at any point within two years of expected high school graduation, variables for enrollment at two-year and four-year institutions are not mutually exclusive measures.

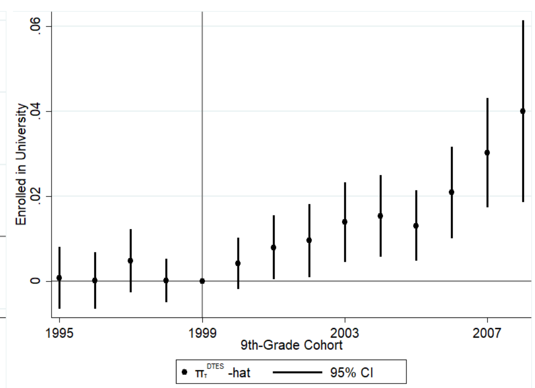
[41] Chetty et al. (2011) find that assignment to small classes (averaging 15 students, rather than 22) for grades K-3 through the Tennessee STAR experiment caused a 1.8 percentage-point increase in college enrollment. Castleman and Long (2016) find that a $1,995 (2020$) increase in need-based aid in Florida caused a 3.2 percentage-point increase in enrollment at public four-year universities. Assuming linearity, my estimated effect on four-year enrollment of 1.6 percentage-points scaled by their coefficient corresponds to a $998 increase in aid.



(a) College Enrollment



(b) CTC Enrollment



(c) University Enrollment

Figure 1.5: Effects of the Labor Demand Shock on College Enrollment

Notes: These figures present estimates of the effect of exposure to negative labor demand shocks during youth and adolescence on college enrollment. Estimates reflect coefficients from two-step event study regressions that compare the difference in outcomes between students from counties that were more or less exposed to negative labor demand shocks across cohorts relative to the continuation of a linear pre-trend (Goodman-Bacon, 2021c). The outcomes are (a) an indicator for enrolling at any public two- or four-year college or university in Texas within two years of expected high school graduation and separate indicators for enrolling at (b) a public two-year community or technical college (CTC) and (c) a public four-year university. The specification controls for county and year fixed effects, student demographics, 1990 county characteristics flexibly interacted with year dummies, county exposure to the fracking boom, and other changes to US trade policy. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable.

ages 16 through 22 is positive, and the nearly 2 percentage-point increases at each age from 18 through 20 are statistically distinguishable from zero at 95% confidence. These estimates suggest that the shock led to sustained increases in college enrollment. Moreover, I find only small and statistically insignificant negative coefficients at older ages, suggesting that students did not merely adjust the timing of college enrollment.

Second, I estimate the effects of exposure to the labor demand shock on cumulative college attainment

Table 1.6: Effects of the Labor Demand Shock on College Enrollment and Degree Receipt

	(1) Enrolled	(2) Enrolled at CTC	(3) Enrolled at Uni	(4) Total Semesters	(5) Certificate	(6) AA/AS	(7) BA/BS
De-trended DD	0.018***	0.012***	0.016***	0.203***	-0.001	-0.001	0.011***
	(0.005)	(0.005)	(0.004)	(0.067)	(0.001)	(0.002)	(0.002)
Percent Change	4.2	3.6	9.5	4.7	-3.1	-3.2	8.3
Pre-period Mean	0.423	0.328	0.173	4.334	0.018	0.043	0.127
N	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707

Standard errors clustered by county in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents estimates of the effects of exposure to negative local labor demand shocks on college enrollment and degree receipt using individual-level data from the University of Houston Education Research Data. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. The outcome in column (1) is an indicator for enrolling in a public two- or four-year college or university within two years of expected HS graduation. The outcomes in columns (2) and (3) are indicators for enrollment at a two-year community or technical college and a four-year university within two years of expected HS graduation, respectively; these variables are not defined to be mutually exclusive (i.e., for a student that enrolls at both a two-year and a four-year institution within two years of expected HS graduation, both indicators will populate as 1). The outcome in column (4) is the total number of semesters a student enrolled in by age 25. The outcomes in columns (5) through (7) are indicator variables for receiving a certificate, associate's degree, and bachelor's degree by age 25, respectively. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

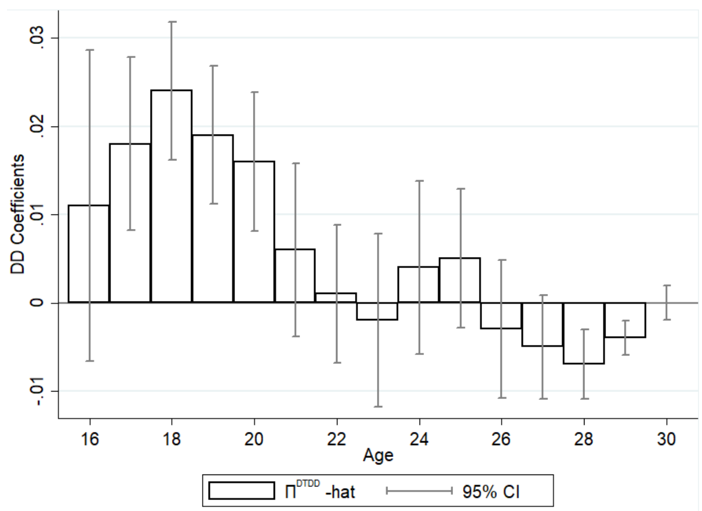


Figure 1.6: Effects of the Labor Demand Shock on College Enrollment by Age

Notes: This figure presents estimates of the effect of exposure to negative labor demand shocks during youth and adolescence on college enrollment (including dual enrollment while in high school) from ages 16-30. Each coefficient represents the estimate from a separate de-trended difference-in-differences regression that compares changes in the differences between outcomes measured at that specific age of students from more- and less-exposed counties among cohorts that reach ninth grade after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c) . The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, county exposure to the fracking boom, and other changes to U.S. trade policy. The outcome in each regression is an indicator variable for being enrolled in a public two- or four-year college in Texas (including dual enrollment while in high school) during that specific age. Students are assigned to the county where they first appeared in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable.

measures and degree receipt in columns (4) through (7) of Table 1.6. Estimates indicate that the shock significantly increased cumulative semesters of college enrollment by age 25 by 5% percent. Affected students were 1.1 percentage-points (8%) more likely to earn a bachelor's degree by age 25, but effects on associate's degree or certificate receipt are small and statistically insignificant. The estimated increase in bachelor's degree receipt may reflect receipt by both marginal enrollees and by students that would have enrolled but left college without a degree in the absence of the shock. Existing research on the persistence toward degree receipt of marginal four-year enrollees suggests that the former group may make a sizable contribution to the estimate (Goodman et al., 2017).

The labor demand shock's incidence on manufacturing may have incentivized students to adjust major choices in similar manners to how results in Section 1.5.1 suggest they shifted course selections while in high school. Using linked college and workforce records for pre-shock graduates, I define major-level exposure to the labor demand shock as the employment-weighted average tariff gap of industries employing graduates from a given field and credential level (i.e., AA/AS vs. BA/BS) in Texas. Table 1.7 presents estimates of the effects of shock exposure on enrollment across fields grouped into quartiles by exposure to the labor demand shock.[42] Consistent with students observing changes to earnings premia across majors, I find statistically significant estimates of enrollment increases in less-exposed majors in Panel A of Table 1.7. Students from more-exposed counties were 55% and 54% more likely to enroll in community college and choose a major in the least-exposed and second-least-exposed quartiles, respectively. I find imprecisely estimated enrollment declines of 5% and 18% for majors the second-most-exposed and most-exposed quartiles, further (suggestively) supporting that community college students substituted from more- to less-exposed majors.

Panel B of Table 1.7 presents the corresponding estimates of effects of local shocks on enrollment across more- and less-exposed majors at four-year universities. Estimates indicate that enrollment increases at universities occurred fairly evenly across majors. The lack of substitution in major choices by university students suggests that the local shocks did not differentially affect the earnings premia for bachelor's degrees across fields of study, consistent with both the more direct link between community college field of study and employment opportunities and evidence that the China shock disproportionately affected earnings for less-educated workers in Section 1.4.2 and other literature (Autor et al., 2014).

1.5.3 Human Capital Adjustments Protect Against Scarring Effects

The above results indicate that in Texas, exposure to the China shock caused students to acquire both more and better-fitting human capital in high school and college. I next assess whether these adjustments to human

[42] Appendix 1.9.2.3 shows the two-digit Classification of Instructional Programs codes in each exposure quartile and Appendix Table 1.17 presents results using major-groupings according to broad categories defined by Foote and Grosz (2020).

Table 1.7: Intensive-Margin Adjustments: Field of Study by Exposure to the China Shock

	(1) Q1	(2) Q2	(3) Q3	(4) Q4
A. Two-Year Colleges				
De-trended DD	0.017***	0.004***	-0.003	-0.005
	(0.003)	(0.001)	(0.010)	(0.003)
Percent Change	55.0	53.9	-4.7	-18.4
Pre-period Mean	0.031	0.007	0.071	0.024
N	3,678,707	3,678,707	3,678,707	3,678,707
B. Four-Year Universities				
De-trended DD	0.005***	0.002***	0.008***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Percent Change	16.9	10.4	17.1	4.6
Pre-period Mean	0.028	0.016	0.044	0.040
N	3,678,707	3,678,707	3,678,707	3,678,707

Standard errors clustered by county in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents estimates of the effects of exposure to negative local labor demand shocks on college enrollment by field of study where majors are grouped according to field-specific shock exposure. I define shock exposure for each major (two-digit CIP code) as the employment-weighted average tariff gap of industries employing recent graduates from that major in the pre-period. I define these field-specific exposure measures separately for two-year and four-year graduates. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. Panel A presents estimates for two-year enrollment outcomes and Panel B presents estimates for four-year enrollment outcomes. Outcomes are indicators for enrollment in fields grouped by field-specific shock quartile. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

capital investments helped protect against lasting declines in earning, known as "scarring effects" (Schwandt and von Wachter, 2019; Oreopoulos et al., 2012; Kahn, 2010). Examining effects of exposure to the shock on earnings is complicated by the perfect collinearity between ninth-grade cohort, age at which earnings are observed, and year of observed earnings. I estimate de-trended event study and difference-in-differences specifications that hold constant the *age* at which earnings are observed. More-exposed students' increased attachment to the education system in their early twenties may negatively bias earnings estimates from specifications that instead hold calendar *year* constant, which would measure earnings at younger ages for cohorts entering high school in the post-period than for those entering high school in the pre-period. On the other hand, estimates from specifications that hold age constant would be positively biased if the effects of PNTR on exposed local labor markets dissipated as time passed. Estimates in Section 1.4.2 – along with the results of Autor et al. (2021) – indicate that there was no such recovery.[43]

Individuals that made key educational decisions before the onset of the China shock may have entered the labor market with skills that were no longer demanded by employers and, thus, have experienced lasting declines in earnings (Liu et al., 2016). Consistent with this notion, individuals from more-exposed counties that entered high school prior to the shock earned \$1,761 (8%) less than their counterparts from less-exposed counties at age 30. However, de-trended difference-in-differences estimates in Table 1.8 show that exposed students who were young enough to adjust their human capital investments experienced statistically significant earnings gains large enough to erase 90% of this gap, despite the persistence of the decline in local labor demand in more-exposed counties. This suggests that human capital adjustments nearly fully buffered against the scarring effects of the shock on earnings that were experienced by older cohorts whom had already made critical educational decisions prior to the onset of the shock. Table 1.8 shows that both extensive-margin (Columns 2 and 3) increases in employment and intensive-margin increases in earnings conditional on employment (Column 4) contributed to these earnings gains.

The scarring effects of negative labor demand shocks are noteworthy for their persistence (Kahn, 2010). I examine whether human capital adjustments provided lasting protection against scarring effects by estimating the effects of exposure to local shocks during youth and adolescence on earnings at each age from 16 through 30 and present results in Figure 1.7. The estimates suggest that adjustments caused sustained earnings increases in their late 20s averaging over \$1,158 each year – despite the continued persistence of overall earnings declines in exposed counties shown in Figure 1.2. Each coefficient from ages 26 to 30 are all statistically distinguishable from zero with at least 95% confidence, and their magnitudes all correspond to substantial reductions in the earnings losses the shock caused for older cohorts.

[43]If anything, Figure 1.2 suggests that labor demand continued to worsen in more-exposed relative to less-exposed counties as time passed from the initial shock. This would bias against finding positive effects of human capital adjustments on earnings.

Table 1.8: Effects of Exposure to the Labor Demand Shock During K-12 on Earnings at 30

	(1) Share of Qtrs Employed	(2) Earnings (2020$)	(3) Employed Every Qtr	(4) Conditional Earnings
De-trended DD	0.015**	1,579***	0.016**	1,802***
	(0.008)	(410)	(0.007)	(267)
Percent Change	2.8	7.3	3.7	4.0
Pre-period Mean	0.513	21,705	0.445	45,561
N	2,135,226	2,135,226	2,135,226	1,074,826

DoF-adjusted standard errors in parentheses are clustered at the county level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effects of exposure to negative local labor demand shocks during youth and adolescence on employment and earnings outcomes at age 30 using individual-level linked data from the University of Houston Education Research Center. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. The outcome in column (1) is the share of quarters at age 30 an individual was employed in Texas. The outcome in column (2) is annual earnings (2020$) at age 30. The outcome in column (3) is an indicator for being employed in every quarter at age 30 in Texas. The outcome in column (4) is earnings (2020$) at age 30, conditional on being employed in Texas in each quarter that year. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

1.5.4 Adjustments by Vulnerable Subgroups

The results in sections 1.5.1 through 1.5.3 indicate that the average student exposed to the China shock in Texas adjusted their human capital investments along extensive and intensive margins and experienced subsequent relative earnings gains. However, not all students may have been able to make these critical adjustments. Therefore, I explicitly examine effects of the shock on particularly vulnerable subgroups of students: male students, students from low-income households, and racial and ethnic minorities.

Male dominance of the manufacturing workforce made men especially susceptible to job displacement from the China shock (Autor et al., 2019).[44] Over one-fifth of employed prime-aged males in Texas worked in manufacturing in 1995, and the salience of manufacturing as a viable career path for males in particular prior to the establishment of PNTR was reflected in the educational choices of male students from areas of Texas that were later exposed to larger shocks. Male students in high-exposure counties were 4 times more likely to complete manufacturing-based electives in high school, 11 times more likely to enroll in manufacturing-based programs at community colleges, and 9 percentage-points less likely to pursue a college education than female students. Thus, human capital adjustments may have been particularly important for male students to buffer against scarring.

[44] Autor et al. (2019) find that the China shock caused larger declines in earnings and employment of men in exposed local labor markets of those of women.

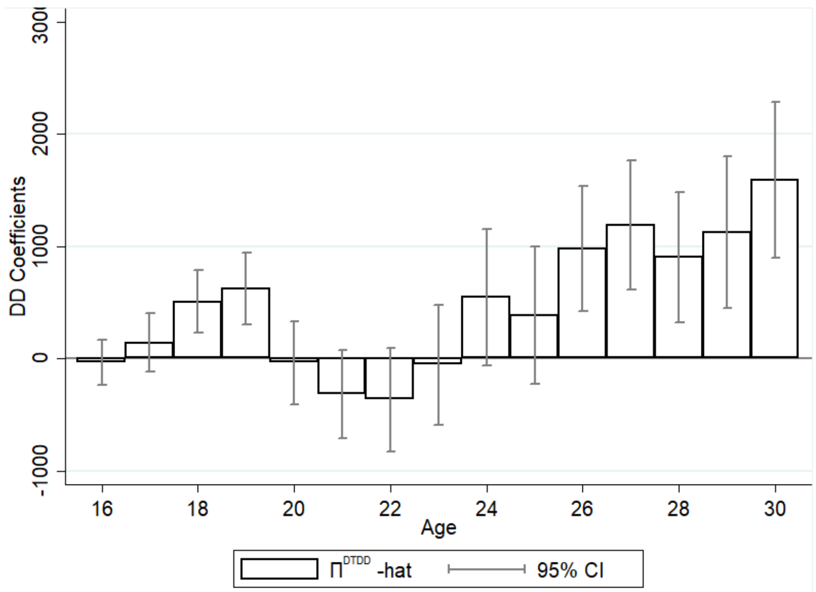


Figure 1.7: Effects of the Labor Demand Shock on Earnings by Age

Notes: This figure presents estimates of the effect of exposure to negative labor demand shocks during youth and adolescence on earnings from ages 16-30. Each coefficient represents the estimate from a separate de-trended difference-in-differences regression that compares changes in the differences between outcomes measured at that specific age of students from more- and less-exposed counties among cohorts that reach ninth grade after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, county exposure to the fracking boom, and other changes to U.S. trade policy. The outcome in each regression is annual earnings from employment in occupations covered by unemployment insurance in Texas. Students are assigned to the county where they first appeared in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable.

I estimate effects of exposure to the shock on educational attainment and later-life earnings for males and present results in Panel A of Table 1.9. The results suggest that exposure to local shocks caused significant increases in dual-credit and AP or IB course-taking, initial and sustained college enrollment, and bachelor's degree receipt. Moreover, male students made dramatic shifts in their fields of study away from those directly affected by the shock, as evident by a 26% decline in manufacturing-based course-taking in high school and a 44% decline in enrollment in manufacturing or construction programs at two-year colleges. Relative to males that made these decisions before the onset of the shock, more-exposed male students earned 9% more at age 30.

Credit constraints may have prevented students from lower-income households from adjusting to the labor demand shock by enrolling in college.[45] However, such constraints should not have hindered students' ability to make adjustments that do not impose direct monetary costs. I proxy for the presence of binding credit constraints with a student's *pre-shock* eligibility for free-or-reduced price lunch (FRPL) and present

[45] Economists debate whether incomplete lending markets for financing the direct monetary costs (i.e., tuition and fees) of attending college result in "credit constraints" that prevent low-income students from obtaining a college education (e.g., Cameron and Taber, 2004; Lochner and Monge-Naranjo, 2012).

Table 1.9: Heterogeneous Effects of the Shock on Vulnerable Populations

	(1) HS Grad	(2) Total Ind. Courses	(3) Total AP/IB Courses	(4) Total Dual-Credit Courses	(5) Enrolled	(6) Total Semesters	(7) BA/BS	(8) Enrolled in Man./Con.	(9) Earnings
A. Male Students									
De-trended DD	0.001	-0.080***	0.075***	0.490***	0.020***	0.253***	0.009***	-0.007***	2,133***
	(0.007)	(0.018)	(0.021)	(0.088)	(0.005)	(0.054)	(0.002)	(0.001)	(525)
Percent Change	0.2	-25.9	53.4	291.2	5.4	6.9	9.0	-43.6	8.8
Pre-period Mean	0.673	0.308	0.141	0.168	0.380	3.651	0.101	0.016	24,224
B. FRPL-Eligible Students									
De-trended DD	-0.012	-0.073***	0.070***	0.468***	0.000	0.047	0.000	-0.005***	1,643***
	(0.009)	(0.016)	(0.019)	(0.066)	(0.005)	(0.050)	(0.001)	(0.001)	(402)
Percent Change	-2.1	-34.3	63.6	388.0	0.1	1.8	1.0	-64.7	9.9
Pre-period Mean	0.606	0.213	0.109	0.121	0.277	2.617	0.049	0.008	16,624
C. Racial and Ethnic Minorities									
De-trended DD	0.002	-0.075***	0.080***	0.687***	0.012**	0.201***	0.007***	-0.005***	1,750***
	(0.009)	(0.016)	(0.022)	(0.098)	(0.006)	(0.075)	(0.002)	(0.001)	(401)
Percent Change	0.4	-36.0	58.0	466.4	3.6	5.8	9.4	-64.6	9.9
Pre-period Mean	0.650	0.208	0.137	0.147	0.340	3.443	0.077	0.008	17,601

DoF-adjusted standard errors clustered by county in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents estimates of the heterogeneous effects of exposure to negative local labor demand shocks on human capital accumulation for subgroups of students that were particularly vulnerable to the shock: male students (A), students eligible for free-or-reduced-price lunch (B), and racial and ethnic minorities (C). Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. The outcome in column (1) is an indicator for graduating from a public TX high school. The outcomes in columns (2) through (4) are total industrial, AP or IB, and dual-credit courses completed, respectively. The outcome in column (5) is an indicator for enrolling in a public two- or four-year college or university within two years of expected HS graduation. The outcomes in columns (6) and (7) are the total number of semesters a student enrolled in by age 25 and an indicator for earning a bachelor's degree by age 25. The outcome in enrollment at a two-year community or technical college and a four-year university within two years of expected HS graduation, respectively; these variables are not defined to be mutually exclusive (i.e., for a student that enrolls at both a two-year and a four-year institution within two years of expected HS graduation, both indicators will populate as 1). The outcome in column (8) is an indicator for enrolling in a manufacturing or construction program at a community or technical college. The outcome in column (9) is reported earnings in positions covered by unemployment insurance in Texas at age 30. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

estimated effects of exposure to the shock on educational attainment and earnings for plausibly constrained students in Panel C.[46] Consistent with the importance of credit constraints for postsecondary investments in human capital, estimates show that FRPL-eligible students did not respond to local shocks by attending college. However, these students took fewer manufacturing-aligned electives and more AP and dual-credit courses in high school – two methods of acquiring college-level human capital that did not impose additional monetary costs. The estimated significant \$1,643 increase in earnings in column (9) suggests the importance of these adjustments.

Negative labor demand shocks typically cause greater employment losses for Black and Hispanic individuals (Hoynes et al., 2012), and the myriad of factors contributing to historical Black-white and Hispanic-white gaps in educational attainment may have hindered human capital adjustments by minority students. I present estimates of the effects of exposure to larger local shocks on students identifying as racial or ethnic minorities in Panel D. The results suggest these students made similar extensive- and intensive-margins as did the student population at large and that these adjustments yielded substantial labor market benefits.

1.6 Mechanisms

In this section, I provide additional context to the main results by examining mechanisms underlying student adjustments.

1.6.1 Did Increased Educational Attainment Reflect Student Learning?

Increases in human capital investment in response to the shock and corresponding earnings gains could represent returns to additional human capital, signaling, or a combination of the two (Spence, 1973). Thus, I examine the effects of local shocks on student learning using data on standardized tests in math scores and reading administered in 8th grade. Estimates in Table 1.10 suggest that students exposed to local shocks performed better on these tests, although only the estimated 0.096 standard deviation improvement is statistically significant ($p < 0.10$). One interpretation of these estimates is that forward-looking students responded to increases in the labor-market *benefits* of higher education by increasing their K-12 academic effort. This behavior is symmetric to existing evidence of the effects of decreases in the *costs* of higher education on academic outcomes in high school (Laajaj et al., 2022; Londoño-Vélez et al., 2020; Bartik and Lachowska, 2014).

[46]Students from households with income below 130% of the poverty line are eligible for free school lunches, and those from households with income between 130% and 185% of the poverty line are eligible for reduced-price lunches. Students from households receiving benefits from means-tested federal programs such as Supplemental Nutritional Assistance Program or Temporary Assistance for Needy Families are automatically eligible for free lunch.

Table 1.10: Effects of the Labor Demand Shock on Standardized Test Scores

	(1) Combined	(2) Math	(3) Reading
De-trended DD	0.144	0.096*	0.048
	(0.112)	(0.054)	(0.060)
N	3,678,707	3,678,707	3,678,707

DoF-adjusted standard errors clustered by county in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents estimates of the effects of exposure to local shocks on 8th-grade test scores, using individual-level data from the University of Houston Education Research Data. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c) The specification controls for county and year fixed effects, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. The outcome variables are combined, math, and reading test stores on exams administered in 8th grade. All test scores are standardized to have a mean of 0 and standard deviation of 1.

1.6.2 Did K-12 and College Supply-Side Responses Help or Hinder Adjustments?

Intensive-margin adjustments across fields of study in both high school and college can be aided or constrained by the responsiveness of programmatic offerings to changes in demand, and students may not be able to access high-demand courses and majors if high schools and colleges have capacity constraints (Grosz, 2022; Grosz et al., 2022). I examine the number of courses and unique programs offered in high school and college, respectively.[47]

I use TEA data to define the average number of course offerings by vocational field across high schools in each county and use IPEDS data to define the number of unique programs (defined by 6-digit CIP codes) within broader major categories (defined by 2-digit CIP codes) across a county's community and technical colleges. Tables 1.18 and 1.19 present estimates of the effects of local shocks on course and major offerings across fields in local high schools and two-year colleges, respectively. I find no statistically significant evidence of "supply-side" adjustments of course or program offerings, although I cannot rule out economically significant effects in either direction.

1.6.3 Did Students Move Across Geographies or Industries?

In addition to adjusting their human capital investments – and in many cases, complementing such adjustments, individuals exposed to local shocks during youth may have moved across geographies, industries, or occupations as adults as manners of adjustment. I first examine whether students exposed to large local shocks migrated to less-exposed counties after high school. Using the location of a student's college or university as a proxy for their adult residence, I estimate effects of greater exposure to the shock on attending college in

[47] An ideal test would estimate effects on the number of *seats* offered in a given course or field of study, but I cannot observe course capacity in my datasets.

any high-exposure (low-exposure) county and on attending college in a high-exposure (low-exposure) county other than where a student attended K-12.[48] Estimates in Table 1.11 suggest that students exposed to larger shocks did not adjust to the shock by moving to less-exposed counties. Greater exposure to the shock significantly increased the likelihood students enrolled in colleges in high-exposure counties without affecting the likelihood they did so in low-exposure counties. Although students exposed to larger shocks were significantly more likely to enroll in colleges outside of their home county, they did so in other high-exposure counties.

Table 1.11: Students Did not Migrate to Less-Exposed Counties for College

	(1) Enrolled in High-Exp. County	(2) Enrolled in Low-Exp. County	(3) Moved for College	(4) Moved to High-Exp. for College	(5) Moved to Low-Exp. for College
De-trended DD	0.017**	-0.001	0.014*	0.013	0.001
	(0.008)	(0.008)	(0.007)	(0.008)	(0.004)
Percent Change	3.8	-1.1	4.5	6.4	1.0
Pre-period Mean	0.439	0.134	0.305	0.198	0.107
N	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707

Standard errors clustered by county in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents estimates of the effects of exposure to negative local labor demand shocks on migration patterns, proxying for a student's adult residence with the location of the college or university they attended. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. The outcome in column (1) and (2) are indicators for enrolling in a public two- or four-year college or university within two years of expected HS graduation in a county with above-median and below-median exposure to PNTR, respectively. The outcome in column (3) is an indicator for enrolling in a college or university outside of the county where a student attended K-12. The outcomes in columns (4) and (5) are indicator variables for enrolling in a college or university in a high-exposure (low-exposure) county other than where a student attended K-12. In all outcomes, students are assigned the location of the modal postsecondary institution they attended. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

I next estimate the effects of exposure to the local shocks on later-life employment across groupings of two-digit NAICS industry codes (Table 1.12). Exposure to shocks during K-12 significantly increased the likelihood that students worked in the manufacturing; construction and transportation; oil and gas; finance, insurance and real estate services; and information services; and health services sectors at age 30.

Due to data limitations, assessing whether exposure to local shocks during youth and the ensuing human capital adjustments affected occupation choiceis beyond the scope of this paper.[49]

[48]Because the UHERC workforce data does not include information on county of residence or employment, I cannot construct employment-based migration measures.

[49]The UHERC dataset includes no information on worker occupation.

Table 1.12: Effects of Exposure to the Labor Demand Shock on Later-Life Employment by Industry

	(1) Man.	(2) Cost./Trans.	(3) Oil and Gas	(4) Retail	(5) Food and Accomm.
De-trended DD	0.004***	0.002**	0.002***	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Percent Change	11.7	5.1	20.8	-0.3	-0.0
Pre-period Mean	0.033	0.042	0.012	0.059	0.036
N	2,135,226	2,135,226	2,135,226	2,135,226	2,135,226
	(6) FIRE	(7) Prof. Services	(8) Information	(9) Admin. Services	(10) Health
De-trended DD	0.004***	0.001	0.002***	0.000	0.003*
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Percent Change	10.3	2.7	15.7	0.9	3.5
Pre-period Mean	0.040	0.033	0.010	0.048	0.075
N	2,135,226	2,135,226	2,135,226	2,135,226	2,135,226

DoF-adjusted standard errors in parentheses are clustered at the county level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of exposure to negative local labor demand shocks during youth and adolescence on employment by industry at age 30 using individual-level linked data from the University of Houston Education Research Center. Industries are categorized by two-digit NAICS codes. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

1.7 Robustness

I test the robustness of my main results to alternative specifications and variable definitions. My primary specification controls for student demographics, 1990 county characteristics interacted with cohort dummies, shale and natural gas presence interacted with cohort dummies, and exposure to other changes to U.S. trade policy.[50] I show that my main results are robust to excluding these controls and to sequentially adding back control variables (Table 1.22). Column (1) corresponds to the preferred specification utilized throughout the paper. Across five specifications, estimates of the increase in college enrollment in Panel A range from 2.9% to 4.2%, and all are statistically significant ($p < 0.10$). Estimates of the increase in earnings at age 30 vary from $1,401 to $1,593 and are all significant ($p < 0.01$).

The China shock was not the only labor demand shock to occur during my sample period, and correlations in the incidence of shocks across counties could confound my estimations and violate the required assumption that differences in outcomes across high-exposure and low-exposure counties would continue to evolve along existing linear trends if both experienced low treatment dosages. In columns (6) through (8) of Table 1.22, I sequentially add controls for county-level exposure to the 2000s housing boom and bust, the 2007-2008 financial crisis, and the 2000 dot-com bubble crash – three labor demand shocks which prior research has shown affected educational attainment (Charles et al., 2018; Weinstein, 2022).[51] Across the three additional specifications, the estimated increase in college enrollment ranges from 3.0% to 5.1%, and the estimated increase in earnings ranges from $1,604 to $1,724. All six estimates are statistically significant ($p < 0.05$).

In addition to including covariates, my primary specification imposes a linear trend in the differences in outcomes across cohorts of students reaching ninth grade prior to the onset of local shocks to account for existing differential secular trends in economic outcomes across high-exposure and low-exposure counties. Causal identification with this specification requires the assumption that differences in outcomes across treatment and control counties would continue to evolve along existing linear trends if both groups of counties experienced only small labor demand shocks. I test for robustness of my main results to relaxing this assumption in two exercises.

First, I test for sensitivity to relaxing the assumption that differential linear trends between high-exposure and low-exposure counties would have continued for the *entire* sample period if they were not exposed to large vs. small shocks. It may instead be reasonable to believe that such differential trends would have only continued for a set period of time. Thus, I estimate specifications that only impose the continuation of linear

[50] These covariates all follow Pierce and Schott (2020), with the exception of controlling for the Texas-specific confounder of fracking boom.

[51] I follow Charles et al. (2018) and specify the size of a county's housing bubble as the magnitude of the largest structural break from trend in housing prices occurring between 2000 and 2006, using county-level housing price indices from the Federal Housing Finance Agency. I specify differential exposure to the financial crisis as a county's pre-period debt-to-income ratio (Mian et al., 2013) and specify a county's exposure to the dot-com crash as their employment share in "high-technology" industries (Hecker, 2005; Weinstein, 2022). I interact each of these cross-sectional exposure measure with cohort dummies to allow them to flexibly affect outcomes across cohorts.

pre-trends for p periods after the onset of the shock before reverting to parallel trends. A standard assumption of parallel trends corresponds to the case of $p = 0$, while the main results correspond to the case of $p = 8$.[52] Panel A of Table 1.23 shows that the estimated increase in college enrollment is still statistically significant when allowing reversion to parallel trends after as few as two post periods. Panel B shows that the estimated increase in later-life earnings is even robust to assuming parallel trends in the post period.

Second, I test for robustness to relaxing the assumption that differences in outcomes between high-exposure and low-exposure counties continued *exactly* linearly after 2000. Following Rambachan and Roth (2023), I allow for smooth deviations from the continuation of linear trends. Table 1.24 presents 95% confidence sets of the estimated effects of exposure to the labor demand shock on college enrollment and earnings when allowing for deviations from the extrapolated linear pre-trend by an additional m each period. Panel A shows that we can still conclude that shock exposure increased college enrollment if we allow additional deviations of up to 0.06 percentage points each period (approximately 3% of the estimated effect size). Panel B shows that the estimated effect on earnings is not robust to even a $10 additional deviation each period.

I find evidence that students exposed to the China shock were more likely to enroll in public colleges in Texas. This result may not reflect an increase in overall attendance if local shocks caused students to substitute from more-expensive out-of-state or private institutions to cheaper in-state public colleges and universities. I assess this threat with two exercises. Although private college enrollment is not available in the UHERC data for my entire analysis period, I test for effects of local shocks on enrolling at not-for-profit private colleges using a subsample of ninth-grade cohorts for whom I can observe private enrollment.[53] Moreover, if local shocks caused students to substitute enrollment in more-expensive colleges to less-expensive colleges, we would expect such an enrollment shift to manifest across the cost distribution *within* Texas, too. I group Texas public universities by their stated in-state tuition and fees using data from the NCES Integrated Postsecondary Education Data System and estimate the effects of exposure to local shocks on enrollment across gross tuition quartiles. Table 1.25 presents results from both of these exercises. Contrary to a story of substitution, I find that exposed students were more likely to attend colleges in all but the third tuition quartile. Moreover, although the estimated effect on private enrollment is negative, the coefficient is not statistically distinguishable from zero and is an order of magnitude smaller than the estimated effect on public university enrollment found in Table 1.6.

[52] Because earnings at age 30 are only observed through the 2002 ninth-grade cohort, the main earnings results correspond to the case of $p = 3$.

[53] Private school enrollment is only available starting in Fall 2002, corresponding to the on-time freshman year for ninth-graders in 1998. I follow Mountjoy (2022) in taking advantage of high persistence rates at private colleges and backward inducting the enrollment timeline for upperclassmen at private colleges in Fall 2002 in previous semesters, assuming on-time persistence.

1.8 Conclusion

This paper examines how students adjust their human capital investments in response to negative shocks to local labor demand and whether these adjustments successfully buffer against harmful effects on labor market outcomes. I use linked student-level administrative data from Texas and leverage variation in local labor demand generated by exposure to Chinese import competition in a de-trended difference-in-differences design. Consistent with reductions in the opportunity cost of schooling, I find that students exposed to larger shocks were 4% more likely to enroll at a two- or four-year college and 8% more likely to obtain a bachelor's degree. Moreover, exposed students completed more courses eligible for college credits while in high school and shifted away from studying fields closely linked to industries that were more-exposed to import competition in favor of those associated with unaffected sectors in both high school and college, suggesting that they internalized salient changes to lifetime earnings premia across attainment levels and fields of study. Following students into the labor market, I provide evidence that these adjustments provided substantial protection against the sustained decline in local labor demand.

Existing research on adverse labor demand shocks – including the China shock – paints a gloomy picture for exposed local labor markets, which still exhibit elevated nonemployment and reduced income more than a decade after the onset of local shocks (Autor et al., 2021; Hershbein and Stuart, 2023). Negative labor demand shocks scar prime-aged workers and new labor market entrants, whom have already made critical educational decisions (Ruhm, 1991; Kahn, 2010; Autor et al., 2014). My findings provide evidence that when individuals coming of age are able to make considerable adjustments to their human capital investments, they can avoid suffering the same fate.

Concerns over adjustment to adverse labor demand shocks are particularly prevalent when shocks represent "permanent" technological or structural changes, as with the China shock. Programs such as Trade Adjustment Assistance help participating prime-aged workers adjust to such shocks (Hyman, 2022). Although there are no analogous programs designed explicitly to help *youth* adjust to structural changes to labor demand, my findings illustrate that forward-looking students can make substantial human capital adjustments in response to changes in local labor demand and suggest considerable labor market returns to such adjustments. In contrast, nationwide studies of similar shocks find that declines in family income and school funding can counteract and even dominate increased incentives for educational attainment (Burga and Turner, 2022; Stuart, 2022). Comparisons of my results with the literature suggest the type of policies that enable students to make human capital adjustments. By preventing declines in local property values from translating into meaningful reductions in local school spending, K-12 finance equalization systems such as Texas' Robin Hood system likely help to protect students from consequences of local shocks. Future research should explicitly examine the roles of state K-12 finance formulas and other policies – such as need-based financial aid programs – in guarding students against adverse economic shocks. Such research may help guide policymakers in designing methods to help students adjust to current and future manifestations of skill-biased technological change, such as automation, the transition away from fossil fuels, and artificial intelligence.

1.9 Appendix

1.9.1 Additional Results

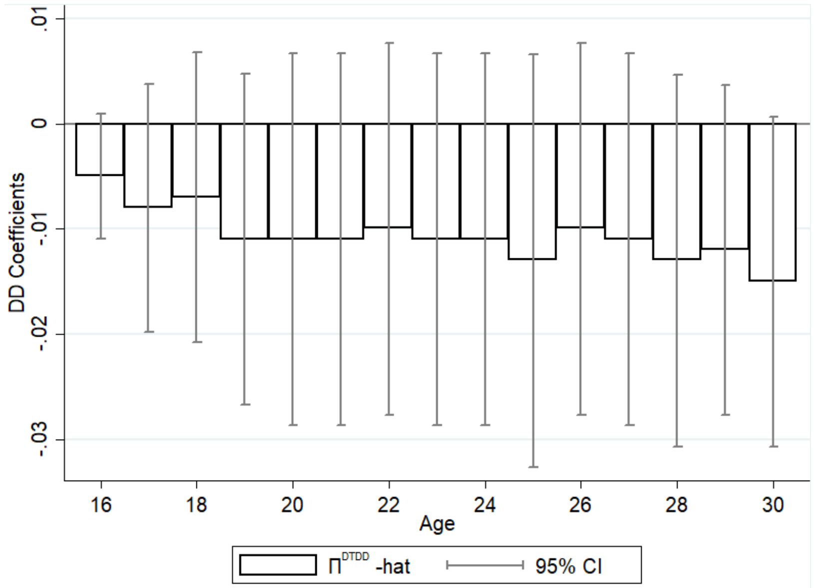


Figure 1.8: No Statistically Significant Evidence of Differential Attrition from the Dataset

Notes: This figure present estimates of the effect of exposure to Chinese import competition on never being observed again in the UHERC dataset from that age through age 30. Each coefficient reflects the estimate of a de-trended difference-in-differences specification that compares changes in outcomes measured at a specific age of students from counties with above- and below-median exposure to local labor market shocks caused by import competition relative to existing differential linear trends (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, 1990 county characteristics flexibly interacted with year dummies, and other changes to US trade policy. Standard errors are clustered by county and adjusted to account for using a regression-adjusted estimator.

Table 1.13: “First-Stage” Effects: Robustness to Alternative Panel Lengths

	(1) Earnings Per Capita	(2) Man. Emp	(3) Total Emp	(4) Emp:Pop
Panel, 1994-2007				
De-trended DD	-1,210***	-894***	-3,392	-0.014**
	(209)	(259)	(2,533)	(0.007)
Panel, 1994-2012				
De-trended DD	-2,194***	-1,443***	-5,370*	-0.030***
	(258)	(462)	(3,158)	(0.008)
Panel, 1994-2016				
De-trended DD	-3,144***	-1,710***	-6,174	-0.039***
	(322)	(570)	(3,960)	(0.008)
Percent Change, 2007	-9.7	-13.8	-8.3	-3.3
Percent Change, 2012		-17.6	-22.3	-13.2
-6.8				
Percent Change, 2016	-25.3	-26.4	-15.2	-8.9
Pre-period Mean	12,432	6,475	40,703	0.436
Number of Counties	254	254	254	254

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of exposure to Chinese import competition on employment-to-population ratios and per capita income, using employment counts from the County Business Patterns Database, population counts by age group from the Survey of Epedemiology and End Results, personal income data from the Bureau of Economic Analysis Regional Economic Accounts dataset, and employment and wage data broken down by age and education levels from the Quarterly Workforce Indicators dataset. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between county-level outcomes in more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). Each row corresponds to an analysis sample with a different panel end-period. The specification controls for county and year fixed effects, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Standard errors are clustered by county and adjusted to account for using a regression-adjusted estimator. The outcome variables in columns (1) and (2) are manufacturing and overall employment in the county. Columns (3) and (4) scale employment and earnings relative to the working-age population. The outcomes and column (5) and (6) are average annual earnings for workers aged 14-24 and for workers without a college degree. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

Table 1.14: Effects of Import Competition on Earnings by Age Group

	(1) 14-24	(2) 25-33	(3) 35-54	(4) 55+
De-trended DD	-1,207***	-1,261***	-429*	-1,036***
	(122)	(191)	(223)	(224)
Percent Change	-18.0	-10.3	-2.8	-7.5
Pre-period Mean	6,722	12,250	15,589	13,843
N	4,286	4,299	4,302	4,299

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of exposure to Chinese import competition on per-capita income by age range, using data on earnings for stable workers broken down by age and education levels from the Quarterly Workforce Indicators dataset and population counts by age group from the Survey of Epedemiology and End Results. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between county-level outcomes in more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Standard errors are clustered by county and adjusted to account for using a regression-adjusted estimator. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

Table 1.15: Effects of Import Competition on Earnings by Education Level

	(1) Less than HS	(2) HS Only	(3) Some College	(4) Bachelor's or More
De-trended DD	-1,261***	-799***	-835***	1,118***
	(153)	(168)	(188)	(378)
perc_change12	-12.5	-6.5	-5.9	5.4
Pre-period Mean	10,101	12,242	14,265	20,819
N	4,298	4,302	4,302	4,302

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of exposure to Chinese import competition on per-worker income by education levels, using data on employment and earnings for stable workers broken down by education levels from the Census Quarterly Workforce Indicators database. The outcome variables are county-level average annual earnings among workers with stable employment among each educational attainment group. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between county-level outcomes in more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Standard errors are clustered by county and adjusted to account for using a regression-adjusted estimator. Standard errors are clustered by county. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

Table 1.16: Effects on Vocational Elective Selections

	(1) All Vocational	(2) Industrial	(3) Technical	(4) Agricultural	(5) Health	(6) Business
De-trended DD	-0.039	-0.047***	-0.023***	0.012*	0.000	0.026
	(0.038)	(0.012)	(0.006)	(0.007)	(0.006)	(0.017)
Percent Change	-2.8	-24.3	-66.2	6.3	0.5	4.0
Pre-period Mean	1.398	0.191	0.035	0.190	0.048	0.631
N	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of exposure to Chinese import competition on vocational elective course completions in high school using individual-level data from the University of Houston Education Research Data. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. The outcome in column (1) is a count of total vocational electives completed in high school and those of the remaining columns correspond to completions of courses in a particular Texas Education Agency category of vocational electives.

Table 1.17: Intensive-Margin Adjustments: Two-Year College Enrollment By Field

	(1) Manufacturing/Construction	(2) IT	(3) Health	(4) Business	(5) Education
Parametric DD	-0.003***	0.002*	0.009***	0.002*	0.000
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Percent Change	-37.4	25.5	56.5	16.0	0.6
Pre-period mean	0.009	0.006	0.015	0.015	0.007
N	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707

Standard errors clustered by county in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents estimates of the effect of exposure to negative local labor demand shocks on two-year college enrollment by field of study using data from the University of Houston Education Research Data. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. Outcomes are indicators for enrolling in a public two-year college in Texas and majoring in a field of study belonging to particular categories defined by Foote and Grosz (2020) based on Classification of Instructional Programs codes. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

Table 1.18: Did Local High Schools Adjust Course Offerings?

	(1) AP/IB	(2) All Vocational	(3) Industrial	(4) Technical	(5) Agricultural	(6) Health	(7) Business
Parametric DD	0.7206	0.1275	-0.9955	0.1092	0.4573	0.1118	0.4048
	(0.5795)	(1.9064)	(1.0153)	(0.0853)	(0.5444)	(0.1312)	(0.6887)
Percent Change	15.5	0.5	-24.8	109.9	5.0	17.3	5.7
Pre-period mean	4.6	26.4	4.0	0.1	8.8	0.7	7.1
	3744	3744	3744	3744	3744	3744	3744

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of exposure to Chinese import competition on course offerings at local high schools. I define average categorical counts of courses across all public high schools in each county. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between county-level outcomes in more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences. The specification controls for county and year fixed effects, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Standard errors are clustered by county.

Table 1.19: Did Local Two-Year Colleges Adjust Program Offerings?

	(1) Q1	(2) Q2	(3) Q3	(4) Q4
De-trended DD	-0.53	-0.94	-0.79	1.47
	(1.39)	(0.69)	(0.83)	(0.91)
Percent Change	-6.1	-26.3	-7.7	28.3
Pre-period Mean	8.6	3.6	10.3	5.2
N	1,313	1,285	1,349	1,340

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of exposure to Chinese import competition on programmatic offerings (six-digit CIP codes) at two-year public colleges by the major's exposure to the shock. I define shock exposure for each major (two-digit CIP code) as the employment-weighted average tariff gap of industries employing recent graduates from that major in the pre-period. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between county-level outcomes in more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Standard errors are clustered by county and adjusted to account for using a regression-adjusted estimator. The outcome variables in each column are the number of unique six-digit CIP code programs offered in a county within that quartile of exposure. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

Table 1.20: Heterogeneous Effects on College Enrollment by Student Demographics

	(1) White	(2) Black	(3) Hispanic	(4) Male	(5) Female	(6) ELL	(7) Non-ELL
De-trended DD	0.010*	0.011	0.032***	0.027***	0.021***	-0.017**	0.020***
	(0.005)	(0.008)	(0.008)	(0.005)	(0.006)	(0.008)	(0.006)
Percent Change	2.0	3.3	9.8	7.1	4.5	-9.1	5.6
Pre-period Mean	0.500	0.336	0.324	0.380	0.467	0.187	0.443

Standard errors clustered by county in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents estimates of the heterogeneous effects of exposure to negative local labor demand shocks on college enrollment by student demographics, using individual-level data from the University of Houston Education Research Data. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. The outcome is an indicator variable for enrolling in a public two- or four-year college in Texas within two years of expected HS graduation. Columns denote the subgroup composing estimating samples: (1) white students, (2) Black students, (3) Hispanic students, (4) male students, (5) female students, (6) English-Language Learners, (7) non-English-Language-Learners. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

Table 1.21: Heterogeneous Effects on Earnings by Student Demographics

	(1) White	(2) Black	(3) Hispanic	(4) Male	(5) Female	(6) ELL	(7) Non-ELL
De-trended DD	1,182**	1,964***	1,883***	2,198***	1,080***	347	1,594***
	(473)	(510)	(398)	(538)	(339)	(545)	(452)
Percent Change	4.6	12.4	10.6	9.1	5.7	2.9	5.7
Pre-period Mean	25,727	15,827	17,784	24,224	19,045	11,875	22,609
N	522,097	138,647	387,744	550,684	524,142	132,000	942,826

Standard errors clustered by county in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents estimates of the heterogeneous effects of exposure to negative local labor demand shocks on earnings at age 30 by student demographics, using individual-level data from the University of Houston Education Research Data. Estimates reflect coefficients from a parametric difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences. The specification controls for county and year fixed effects, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county. The outcome is earnings (2020$) at age 30. Columns denote the subgroup composing estimating samples: (1) white students, (2) Black students, (3) Hispanic students, (4) male students, (5) female students, (6) English-Language Learners, (7) non-English-Language-Learners. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

Table 1.22: Robustness of Main Results to Sequentially Adding Covariates

	(1) Preferred (Col. (5) + Fracking)	(2) No Controls	(3) + Demos	(4) + County Covariates	(5) + Trade Policies	(6) + Housing Bubble	(7) + Financial Crisis	(8) + Dot-Com Crash
A. College Enrollment								
De-trended DD	0.018***	0.012*	0.012*	0.014***	0.018***	0.021***	0.019***	0.013**
	(0.005)	(0.007)	(0.007)	(0.004)	(0.005)	(0.006)	(0.005)	(0.005)
Percent Change	4.2	2.9	2.9	3.3	4.2	5.1	4.6	3.0
Pre-period Mean	0.423	0.423	0.423	0.423	0.423	0.423	0.423	0.423
N	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707
B. Earnings								
De-trended DD	1,579***	1,431***	1,450***	1,401***	1,593***	1,724***	1,810***	1,604***
	(410)	(352)	(354)	(427)	(411)	(403)	(403)	(360)
Percent Change	7.3	6.6	6.7	6.5	7.3	7.9	8.3	7.4
Pre-period Mean	21,705	21,705	21,705	21,705	21,705	21,705	21,705	21,705
N	1,074,826	1,074,826	1,074,826	1,074,826	1,074,826	1,074,826	1,074,826	1,074,826

DoF-adjusted standard errors in parentheses are clustered at the county level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of exposure to negative local labor demand shocks on college enrollment and later-life earnings, using individual-level data from the University of Houston Education Research Data. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences. The outcome variable in Panel A is an indicator variable for enrolling in a public two- or four-year college in Texas within two years of expected HS graduation. The outcome variable in Panel B is earnings at age 30, conditional on being employed in all four quarters in Texas. Column (1) controls for student-level demographics, 1990 county characteristics interacted with a post dummy, exposure to changes in other trade policies, and exposure to the fracking boom and corresponds to the preferred specification throughout the paper. Column (2) includes no additional controls and columns to the right sequentially add control variables. Column (3) controls for student demographics. Column (4) controls for 1990 county characteristics interacted with a post dummy. Column (5) controls for changes to other trade policies. Column (6) controls for the housing boom. Column (7) controls for exposure to the financial crash. Column (8) controls for exposure to the dot-com bubble. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

Table 1.23: Robustness to Reversions to Parallel Trends

	(1) Parallel in Post	(2) Linear through 1	(3) Linear through 2	(4) Linear through 3	(5) Linear through 4	(6) Linear through 5	(7) Linear through 6	(8) Linear through 7	(9) Linear through 8
A. College Enrollment									
De-trended DD	0.005	0.008	0.010*	0.012**	0.014***	0.015***	0.016***	0.017***	0.017***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Percent Change	1.2	1.8	2.4	2.8	3.2	3.6	3.8	4.0	4.1
Pre-period Mean	0.423	0.423	0.423	0.423	0.423	0.423	0.423	0.423	0.423
N	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707	3,678,707
B. Earnings									
De-trended DD	1,113***	1,347***	1,502***	1,579***					
	(410)	(410)	(410)	(410)					
Percent Change	5.1	6.2	6.9	7.3					
Pre-period Mean	21,705	21,705	21,705	21,705					
N	1,074,826	1,074,826	1,074,826	1,074,826					

DoF-adjusted standard errors in parentheses are clustered at the county level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of exposure to negative local labor demand shocks on college enrollment and later-life earnings, using individual-level data from the University of Houston Education Research Data. Estimates reflect coefficients from modifications of a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to a continuation of existing trends that reverts to parallel trends after p periods. The outcome variable in Panel A is an indicator variable for enrolling in a public two- or four-year college in Texas within two years of expected HS graduation. The outcome variable in Panel B is earnings at age 30, conditional on being employed in all four quarters in Texas. Column (1) presents estimates where the first-step partials out only the linear pre-trend and assumes that outcomes revert to trending in parallel in the post period. Columns (2) through (9) partial out present estimates where the first-step partials out a linear pre-trend that extends p periods into the post period before reverting to parallel. Column (9) corresponds to the preferred specification used throughout the main analyses in Panel A. Earnings at age 30 is only observed through the 2002 ninth-grade cohort, so Column (4) correspondes to the preferred specification used throughout the main analyses in Panel (B)

Table 1.24: Robustness to Smooth Deviations from Linear Trends (Rambachan and Roth, 2023)

	(1) Deviation Size (m)	(2) 95% Confidence Set Lower Bound	(3) 95% Confidence Set Upper Bound
A. College Enrollment			
	0.0002	0.008	0.048
	0.0004	0.005	0.053
	0.0006	0.002	0.058
	0.0008	-0.002	0.064
	0.0010	-0.005	0.069
A. Earnings at 30			
	10	-178	3,932
	20	-198	3,965
	30	-217	3,980
	40	-237	4,013
	50	-257	4,027

Notes: This table presents results from tests for the robustness of main estimates to allowing smooth deviations from the imposed continuation of linear trends (Rambachan and Roth, 2023). Estimates reflect robust 95% confidence sets that account for both statistical uncertainty and allow for the slope of the difference in trends between high-exposure and low-exposure students to deviate by an additional m in either direction each period. Inference is based on fixed length confidence intervals (Rambachan and Roth, 2023). original estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset.

Table 1.25: Four-Year Enrollment Responses Across Tuition Quartiles

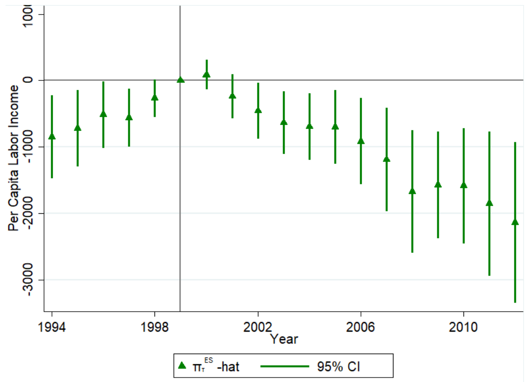
	(1) Q1 Tuition	(2) Q2 Tuition	(3) Q3 Tuition	(4) Q4 Tuition	(5) Private
De-trended DD	0.008***	0.003*	-0.005**	0.008***	-0.002
	(0.002)	(0.002)	(0.003)	(0.002)	(0.001)
Percent Change	57.3	8.3	-18.9	9.0	-7.1
Pre-period Mean	0.014	0.038	0.028	0.094	0.028
N	3,678,707	3,678,707	3,678,707	3,678,707	3,154,442

Standard errors clustered by county in parentheses

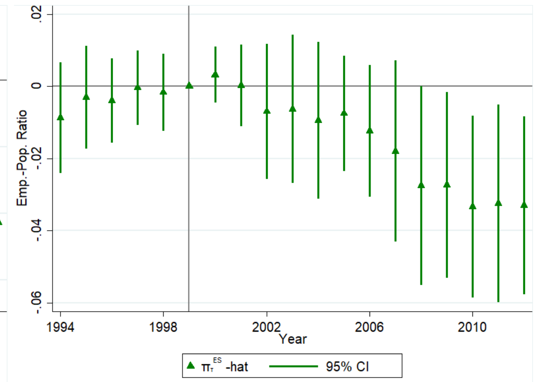
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents estimates of the effect of exposure to negative local labor demand shocks on enrollment at public four-year college enrollment across the distribution of in-state tuition and fees and at private four-year colleges. Estimates reflect coefficients from a de-trended difference-in-differences regression that compares changes in the differences between outcomes of students from more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences (Goodman-Bacon, 2021c). The specification controls for county and year fixed effects, student demographics, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Students are assigned to the county of the first school where they appear in the dataset. Standard errors are clustered by county and adjusted to account for using a regression-adjusted outcome variable. Standard errors are clustered by county. Outcomes in columns (1) through (4) are indicators for enrolling within two years of expected HS graduation at public four-year colleges grouped by quartile of 1999 in-state tuition and fees. The outcome in column (5) is an indicator for enrolling within two years of expected HS graduation at a private college Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

1.9.1.1 Unadjusted Event Studies

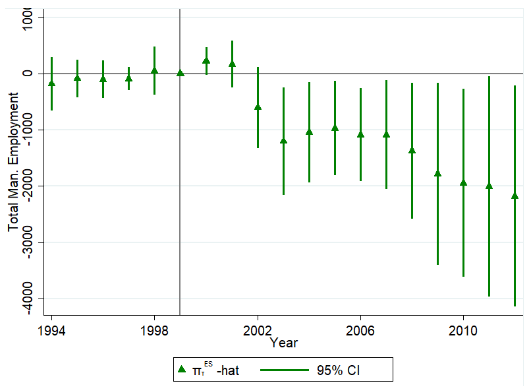


(a) Per-Capita Labor Income

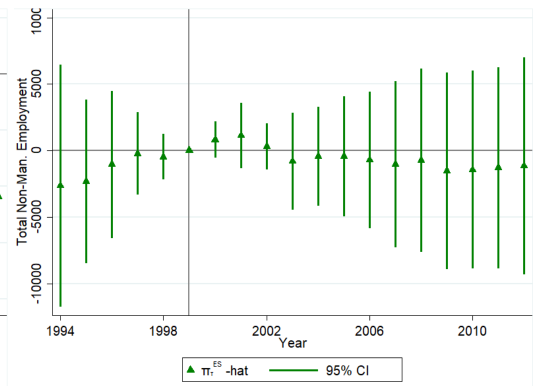


(b) Employment:Population

Figure 1.9: Employment:Population



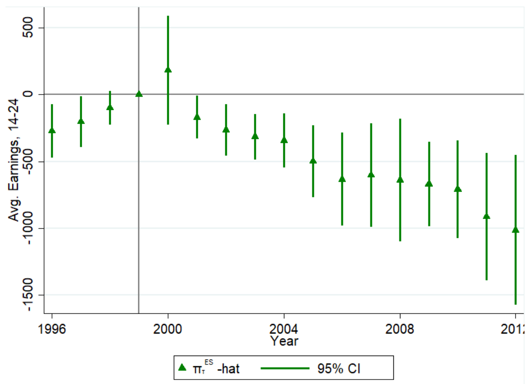
(a) Manufacturing Employment



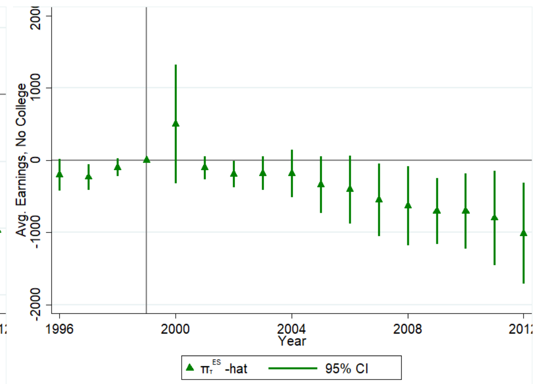
(b) Non-Manufacturing Employment

Figure 1.10: "First-Stage" Effects of Import Competition on Local Labor Markets

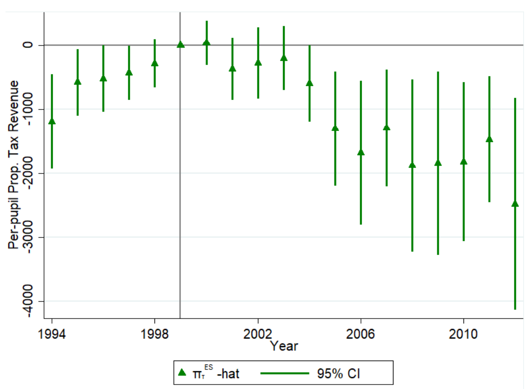
Notes: These figures present estimates of the effect of exposure to Chinese import competition on employment and labor income relative to population using employment counts from the County Business Patterns Database, population counts by age group from the Survey of Epedemiology and End Results, and personal income data from the Bureau of Economic Analysis Regional Economic Accounts dataset. Estimates reflect coefficients from event study regressions (equation (1.1)) that compare the difference in outcomes between counties with above- and below-median exposure to local labor market shocks caused by import competition over time relative to this relationship in 1999, one year prior to the start of treatment. The specification controls for county and year fixed effects, 1990 county characteristics flexibly interacted with year dummies, county exposure to the fracking boom, and other changes to US trade policy. Standard errors are clustered by county.



(a) Earnings, School-Aged Workers



(b) Earnings, Workers Without College Experience

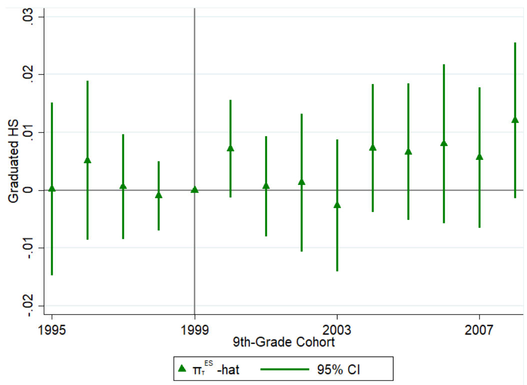


(c) Per-Pupil Property Tax Revenue

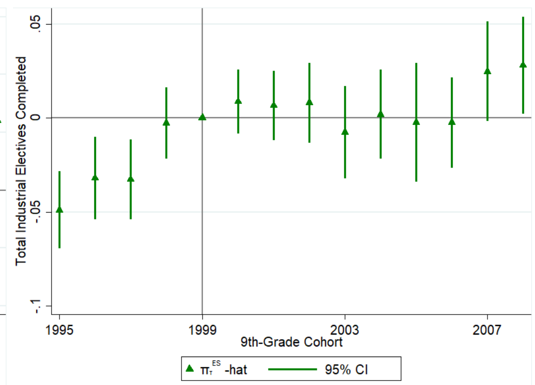
(d) Per-Pupil K-12 Expenditure

Figure 1.11: The Labor Demand Shock Reduced Opportunity Costs But Did Not Affect K-12 Spending

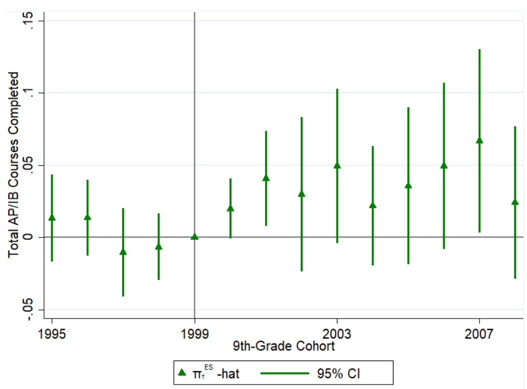
Notes: These figures present estimates of the effect of exposure to Chinese import competitionon opportunity costs and school spending using wage data broken down by age and education levels from the Quarterly Workforce Indicators dataset and K-12 finance data from the National Center for Education Statistics Common Core of Data. Estimates reflect coefficients from event study regressions (equation (1.1)) that compare the difference in outcomes between counties with above- and below-median exposure to local labor market shocks caused by import competition over time relative to this relationship in 1999, one year prior to the start of treatment. The specification controls for county and year fixed effects, 1990 county characteristics flexibly interacted with year dummies, and other changes to US trade policy. Standard errors are clustered by county.



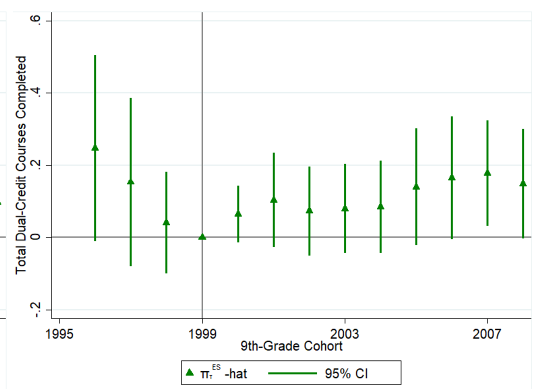
(a) HS Graduation



(b) # of Industrial Electives



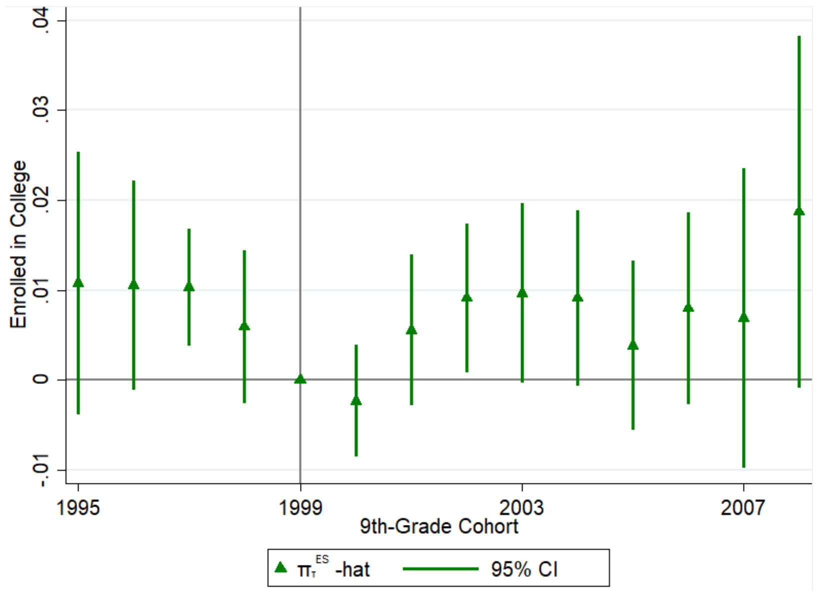
(c) # if AP/IB Courses



(d) Dual Enrollment

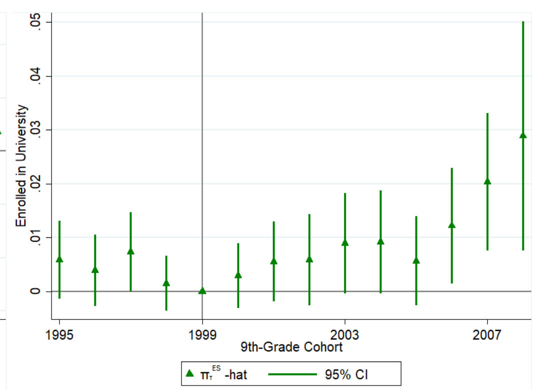
Figure 1.12: Effects of the Labor Demand Shock on Human Capital Accumulation in HS

Notes: These figures present estimates of the effect of exposure to adverse local shocks during youth and adolescence on high school graduation and course selection. Estimates reflect coefficients from event study regressions (equation (1.1)) that compare the difference in outcomes between students from counties that were more or less exposed to negative labor demand shocks across cohorts relative to this relationship for the cohort entering ninth grade in 1999, one year prior to the onset of local shocks. The outcomes are (a) an indicator for graduating high school, counts of the number of (b) manufacturing-aligned electives and (c) AP/IB courses completed, and (d) an indicator variable for enrolling in a dual-credit course at a local college. The specification controls for county and year fixed effects, student demographics, 1990 county characteristics flexibly interacted with year dummies, county exposure to the fracking boom and other changes to US trade policy. Standard errors are clustered by county.



(a) College Enrollment

(b) CTC Enrollment

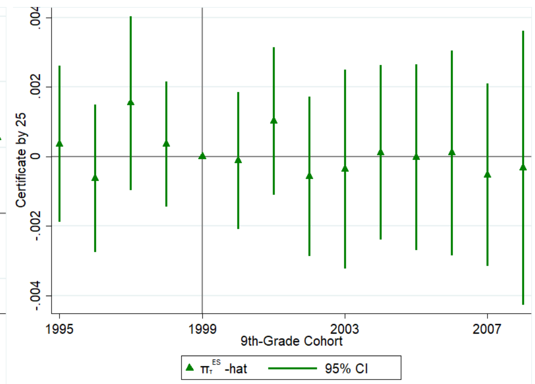


(c) University Enrollment

Figure 1.13: Effects of the Labor Demand Shock on College Enrollment

Notes: These figures present estimates of the effect of exposure to adverse local shocks during youth and adolescence on college enrollment. Estimates reflect coefficients from event study regressions (equation (1.1)) that compare the difference in outcomes between students from counties that were more or less exposed to negative labor demand shocks across cohorts relative to this relationship for the cohort entering ninth grade in 1999, one year prior to the onset of local shocks. The outcomes are (a) an indicator for enrolling at any public two- or four-year college or university in Texas within two years of expected high school graduation and separate indicators for enrolling at (b) a public two-year community or technical college (CTC) and (c) a public four-year university. The specification controls for county and year fixed effects, student demographics, 1990 county characteristics flexibly interacted with year dummies, county exposure to the fracking boom, and other changes to US trade policy. Standard errors are clustered by county.

(a) Total Semesters



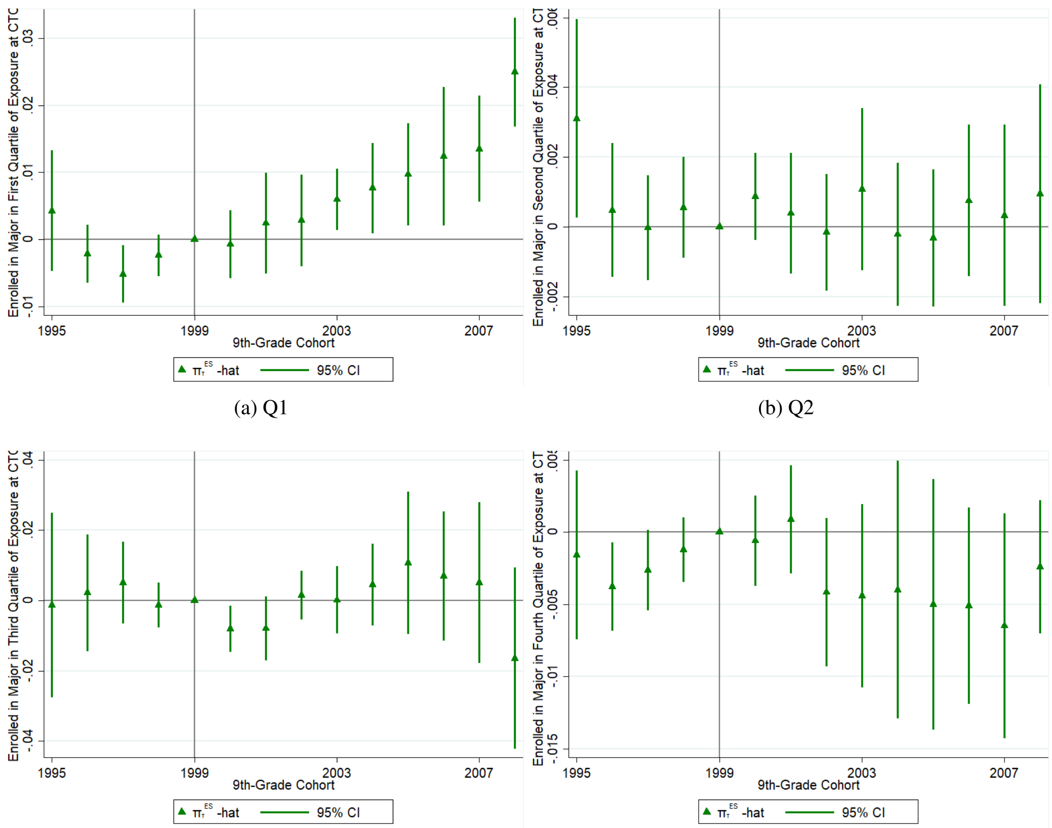
(b) CTC Certificate Receipt

(c) Associate's Degree Receipt

(d) Bachelor's Degree Receipt

Figure 1.14: Effects of the Labor Demand Shock on College Attainment

Notes: These figures present estimates of the effect of exposure to adverse local shocks during youth and adolescence on human capital accumulation. Estimates reflect coefficients from event study regressions (equation (1.1)) that compare the difference in outcomes between students from counties that were more or less exposed to negative labor demand shocks across cohorts relative to this relationship for the cohort entering ninth grade in 1999, one year prior to the onset of local shocks. The outcomes are total semesters enrolled in public colleges and universities in Texas by age 25 (a), receipt of a technical certificate from a two-year college by age 25 (b), associate's degree receipt by 25 (c), and bachelor's degree receipt by 25 (d). Students are assigned to the county where they first appeared attending school prior to the start of the shock. The specification controls for county and year fixed effects, student demographics, 1990 county characteristics flexibly interacted with year dummies, county exposure to fracking, and other changes to US trade policy. Standard errors are clustered by county.

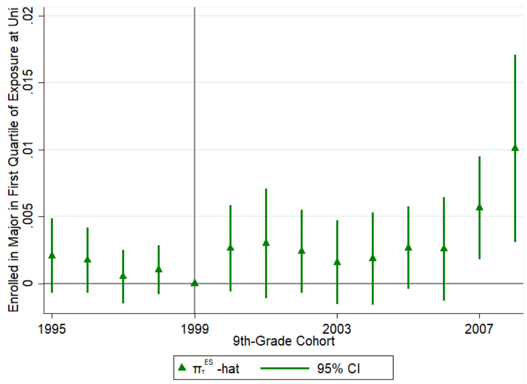


(a) Q1 (b) Q2

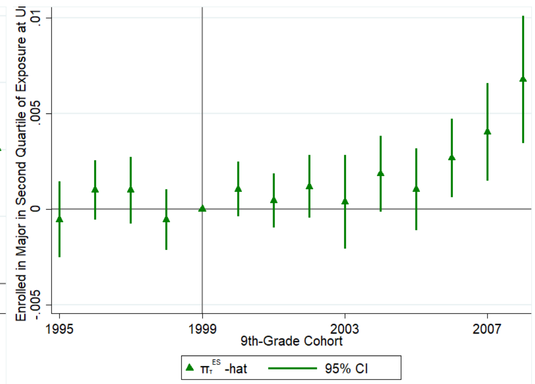
(c) Q3 (d) Q4

Figure 1.15: Effects on Enrollment by Field Exposure Quartile at Two-Year Colleges

Notes: These figures present estimates of the effect of exposure to adverse local shocks during youth and adolescence on enrollment by major at public two-year colleges. I define shock exposure for each major (two-digit CIP code) as the employment-weighted average tariff gap of industries employing recent graduates from that major in the pre-period and bin majors into quartiles. Estimates reflect coefficients from event study regressions (equation (1.1)) that compare the difference in outcomes between students from counties that were more or less exposed to negative labor demand shocks across cohorts relative to this relationship for the cohort entering ninth grade in 1999, one year prior to the onset of local shock. Outcomes are indicator variables for enrollment in a two-year college and selection of a major within that quartile of exposure. Students are assigned to the county where they first appeared attending school prior to the start of the shock. The specification controls for county and year fixed effects, student demographics, 1990 county characteristics flexibly interacted with year dummies, county exposure to fracking, and other changes to US trade policy. Standard errors are clustered by county.

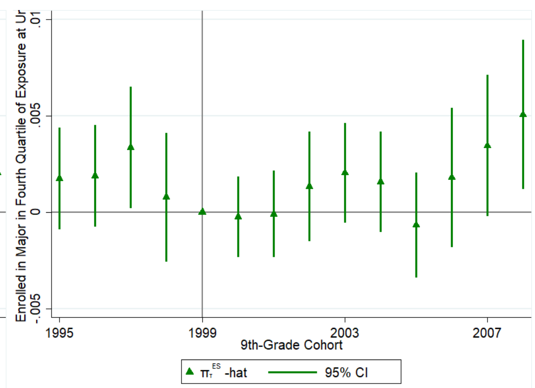


(a) Q1



(b) Q2

(c) Q3



(d) Q4

Figure 1.16: Effects on Enrollment by Field Exposure Quartile at Four-Year Colleges

Notes: These figures present estimates of the effect of exposure to adverse local shocks during youth and adolescence on enrollment by major at public four-year universities. I define shock exposure for each major (two-digit CIP code) as the employment-weighted average tariff gap of industries employing recent graduates from that major in the pre-period and bin majors into quartiles. Estimates reflect coefficients from event study regressions (equation (1.1)) that compare the difference in outcomes between students from counties that were more or less exposed to negative labor demand shocks across cohorts relative to this relationship for the cohort entering ninth grade in 1999, one year prior to the onset of local shock. Outcomes are indicator variables for enrollment in a four-year university and selection of a major within that quartile of exposure Students are assigned to the county where they first appeared attending school prior to the start of the shock. The specification controls for county and year fixed effects, student demographics, 1990 county characteristics flexibly interacted with year dummies, county exposure to fracking, and other changes to US trade policy. Standard errors are clustered by county.

1.9.2 Data

1.9.2.1 Sample Selection

My primary analysis sample consists of 3,678,707 students that attended ninth grade at a public high school in Texas between fall 1995 and 2008. I assign cohort based on a student's first attempt at ninth grade. Because I am interested in adjustment mechanisms that occur during high school, I limit the sample to students that I ever observe starting ninth grade.[54] In order to an adopt an intent-to-treat framework that accounts for endogenous migration, I also limit the sample to students that I observe in the data *before* the onset of the labor demand shock in 2000. This rule produces the 2008 cohort endpoint: on-time ninth-graders in 2008 attended kindergarten in 1999, making them the youngest students I can assign to a county prior to the onset of the shock. Although my dataset starts in 1994, I do not begin my analysis sample until the 1995 ninth-grade cohort, so that I can observe eighth-grade standardized test scores. I drop students participating in special education, because of their unique labor market opportunity set relative to the rest of the sample and the labor demand shock.

Not all students that attend a public high school in Texas remain in the state after exiting K-12. If out-migration from Texas – or, complete detachment from the labor force – are correlated with treatment, then selective attrition from my dataset could introduce non-classicial measurement error that biases estimations of the effects on exposure to local shocks on college enrollment and labor market outcomes in a direction determined by the selection. However, if attrition is uncorrelated with treatment, then it would only represent classical measurement error and bias my results toward a null result. Figure 1.8 presents estimates of de-trended difference-in-difference specifications representing changes in the difference in attrition rates between students from more- and less-exposed counties relative to existing trends. Each coefficient represents an estimate from a separate specification with the outcome variable as never being observed across K-12, postsecondary, or workforce records in Texas from that age through age 30. No coefficient is statistically distinguishable from zero, suggesting attrition only causes classical measurement error. The increase in magnitude of coefficients as the age of definition approaches 30 largely because of a mechanical effect. Because I observe no workers past age 30, "attrition" at age 29 only reflects non-participation in the labor force or education system in Texas for 2 years, as opposed to non-participation for 10 years for attrition defined at age 20. To the extent that human capital adjustments increased labor force opportunities, increases in employment for a given year appear as reduced attrition for later ages, while more weakly relating to earlier defined measures. I drop all students from the analysis sample that attrit at age 17 or earlier and measure employment and earnings outcomes for stayers that exhibit stretches of detachment from the dataset as zeroes when not observed.

[54] I cannot observe comprehensive course completions for students that do not appear in the dataset until tenth grade or later.

1.9.2.2 Other Data Sources

Exposure to PNTR: I define exposure to Chinese import competition following the establishment of Permanent Normal Trade Relations using industry-level and county-level tariff gaps from Pierce and Schott (2016) and Pierce and Schott (2020). Industry-level tariff gaps are defined as the differences between Normal Trade Relations and non-Normal Trade Relations *ad valorem* tariff rates in 1999 for all four-digit Standard Industrial Classification codes (Feenstra et al., 2002). Pierce and Schott (2020) define county-level tariff gaps as the employment-weighted industry-level tariff gaps across all industries present in the county using 1990 industry-level employment counts from the US County Business Patterns database. *Baseline County Characteristics:* My preferred de-trended difference-in-differences specification includes 1990 county characteristics interacted with individual year dummies to flexibly control for confounders related to county economic profiles. I follow Pierce and Schott (2020) in controlling for the percent of the population without any college education, median household income, the foreign-born population share, and the share of employment in manufacturing.[55] The first three measures come from the Census Bureau's 1990 Decennial Census, while the fourth comes from the County Business Patterns database.

Trade Policy Changes: I follow Pierce and Schott (2020) in controlling for four time-varying measures of a county's exposure to changes in trade policy aside from PNTR. First, I include a county's annual labor-share-weighted import tariff rate (i.e., the average tariff rate among goods produced in that county) under Normal Trade Relations, which ensures that the identifying variation does not reflect changes to preferred tariff rates for all countries subject to Normal Trade Relations prior to the establishment of PNTR with China. I also control for a county's exposure to the phasing out of quotas on textile and clothing imports under the Multifiber Agreement during the 1990s and 2000s. The relaxation of quotas occurred over four phases (January 1, 1995, 1998, 2002, and 2005), and quotas on Chinese imports were not affected until after it joined the WTO in 2001. The county-year-level measure of exposure to the quota eliminations weights the quota fill rate (i.e., how binding the quota was during a particular phase) for each industry in a county by the employment share in that particular industry. Finally, upon joining the WTO in 2001, China reduced its import tariff rates and reduced production subsidies, potentially increasing demand for exports of U.S. manufactured goods. I control for individual year dummies interacted with (1) a labor-share-weighted-average of the change in Chinese import tariffs across industries for each county and (2) a labor-share-weighted average of the change in production subsities across industries for each county. Pierce and Schott (2020) compute these measures using product-level data on Chinese import tariffs from Brandt et al. (2017) and subsidies reported in the Chinese National Bureau of Statistics' Annual Report of Industrial Enterprise Statistics.

[55] Unlike (Pierce and Schott, 2020), I do not control for the Veteran population share, a potential confounder specific to their outcome of interest of deaths of despair.

Fracking: Innovation in oil and gas extraction technology in the early 2000s made extraction of previously written-off oil and gas deposits contained in shale formations suddenly economically and technically feasible. The ensuing "fracking boom" across shale-rich regions of Texas increased labor market opportunities for workers with little education, negatively affecting educational attainment (Kovalenko, 2023). I control for exposure to the fracking boom as the per-capita energy potential (in millions of British Thermal Units) of the shale reserves beneath a county's borders (Kovalenko, 2023; Cascio and Narayan, 2022). I use shapefiles of shale plays and maxium estimates of shale oil and gas reserve volumes from the Energy Information Administration (EIA).

Other Labor Demand Shocks: In robustness exercises, I control for a county's exposure to three additional labor demand shocks during my sample period: the 2000s housing boom and bust, the 2007-2008 financial crisis, and the 2000 dot-com bubble crash – three labor demand shocks which prior research have shown affected educational attainment (Charles et al., 2018; Weinstein, 2022). I follow Charles et al. (2018) and specify the size of a county's housing bubble as the magnitude of the largest structural break from trend in housing prices occurring between 2000 and 2006, using county-level housing price indices from the Federal Housing Finance Agency. I specify differential exposure to the financial crisis as a county's pre-period debt-to-income ratio (Mian et al., 2013) using data from the Federal Reserve Enhanced Financial Accounts database. Finally, I specify a county's exposure to the dot-com crash as their employment share in "high-technology" industries (Hecker, 2005; Weinstein, 2022).[56] I interact each of these cross-sectional exposure measure with year dummies to allow them to flexibly affect outcomes across cohorts.

Employment and Earnings: I use county-level measures of employment and earnings outcomes from three data sources. County-level employment counts (overall and by industry) fome from the County Business Patterns Database (Eckert et al., 2021), which imputes missing industry-county-year cells from the Census Bureau County Business Patterns dataset that are suppressed for confidentiality and standardizes consistent industry codes across years. County-level aggregate personal income – broken down by wage and salary income and government transfer income – come from the Bureau of Economic Analysis Regional Economic Accounts dataset. I use samples of both of these datasets starting in 1994, the first year of my student-level panel. I supplement these datasets – which reflect entire county populations – with data on earnings and employment broken down by subgroups of interest (education levels, age groups, race, and gender) from the Census Quarterly Workforce Indicators (QWI) dataset. The QWI reflects the combination

[56]The level-I high-technology industries defined by Hecker (2005) are as follows: pharmaceutical and medicine manufacturing; computer and peripheral equipment manufacturing, communications equipment manufacturing; semiconductor and other electronic component manufacturing; navigational, measuring, electromedical, and control instruments manufacturing; aerospace product and parts manufacturing; software publishers; internet service providers and web search portals; data processing, hosting, and related services; architectural, engineering, and related services; computer systems design and related services; scientific research-and-development services.

of state Unemployment Insurance earnings data and the Quarterly Census of Employment and Wages data with administrative Census Bureau data through the Longitudinal Employer-Household Dynamics program and provides statistics on job flows, employment, and earnings for uniquely detailed geographies, firm, and worker characteristics. Although my UHERC administrative dataset already includes earnings and employment records for all workers in positions covered by Unemployment Insurance in Texas, I can only observe demographics and education levels for workers that attended K-12 or college in Texas during my sample period, making the QWI preferable for first-stage estimations of the effects of the labor demand shock on earnings by age groups and education levels. However, the QWI does not start until 1996, so my analysis period when estimating specifications using QWI-based outcomes differs slightly from that of my primary estimations.

Population: I use annual county-level population counts by age group from the Surveillance, Epeidemiology, and End Results dataset to construct employment rates and per-capita measures of outcomes from the above datasets. I adopt the standard definition of the working-age population as individuals from 15 to 64 and prime-age workers as 25 - 54.

K-12 District Spending: To construct per-pupil measures of categorized K-12 school district revenues and expenditures, I use data from the National Center for Education Statistics Local Education Agency Finance Survey (F-33) and Local Education Agency Universe Survey.

Converting Nominal to Real Measures: I use the Bureau of Labor Statistics Consumer Price Index to convert all nominal measures to 2020 dollars.

1.9.2.3 HS Elective & College Major Categories

Table 1.26: High School Vocational Course Subject Groups

Group	TEA Subject Areas	Example Courses
Agriculture	Agricultural Science (63)	Agricultural Mechanics
Business	Business Education (70); Marketing (65); Office Education (67)	Business Management; Principles of Marketing; Introduction to Computers
Health	Health (81)	Sports Medicine
Industrial	Industrial/Tech Electronics (59); Industrial Arts (60); Trade and Industrial (62)	Digital Electronics; Manufacturing Systems; Intro. to Preceision Metal Manufacturing
Technology	Technology Education (69)	Electricity/Electronics Technology

Notes: This table presents definitions and example courses from categories of high school vocational electives. I group courses belonging to similar Texas Educational Agency Subject Areas to define broad vocational categories. Estimations of the effects of exposure to local shocks on course completions by vocational course category are presented in Table 1.16.

Table 1.27: Quartiles of Major-Level Shock Exposure at Two-Year Colleges

Quartile	CIP Codes	CIP Descriptions
Q1:	02, 13, 19, 22, 25, 31, 32, 34, 36, 43, 44, 45, 49, 51, 54	Agriculture; Education; Home Economics; Legal Studies; Library Science; Recreation Studies; Basic Skills; Health-Related Skills; Leisure; Protective Services; Public Administration; Social Sciences; Transportation; Health Professions; History
Q2:	01, 08, 09, 10, 12, 26, 42	Agricultural Bus. & Prod.; Marketing; Communications; Communications Tech.; Personal Services; Biological Sciences; Psychology
Q3:	16, 24, 52	Foreign Languages; General Studies; Business
Q4:	03, 04, 05, 11, 14, 15, 20, 23, 27, 30, 38, 39, 40, 41, 46, 47, 48, 50	Natural Resources; Architecture; Cultural Studies; Information Sciences; Engineering; Engineering Tech.; Vocational Home Ec.; English; Mathematics; Interdisciplinary; Philosophy; Religious Vocations; Physical Sciences; Science Tech.; Construction Trades; Mechanics and Repairers; Precision Production Trades; Visual & Performing Arts

Notes: This table presents fields of study at two-year colleges grouped into quartiles of major-level exposure. Majors are defined as two-digit Classification of Instructional Programs codes, using NCES crosswalks from 1990 to 2000 and 2010 to make consistent codes. I define major-level shock exposure as the employment-weighted average tariff gap of industries employing recent graduates from that major in the pre-period. Estimations of the effects of exposure to local shocks on enrollment across major quartiles are presented in Table 1.7.

Table 1.28: Quartiles of Major-Level Shock Exposure at Four-Year Universities

Quartile	CIP Codes	CIP Descriptions
Q1:	12, 13, 19, 22, 25, 30, 31, 32, 36, 43, 44, 46, 47, 49, 51, 54	Personal Services; Education; Home Economics; Legal Studies; Library Science; Interdisciplinary Studies; Recreation Studies; Basic Skills; Leisure; Protective Services; Public Administration; Construction Trades; Mechanic & Repair Technologies; Transportation; Health Professions; History
Q2:	02, 16, 26, 42	Agricultural Bus. & Prod.; Marketing; Foreign Languages; Biological Sciences; Psychology
Q3:	01, 03, 04, 05, 09, 23, 24, 27, 38, 45, 50	Agriculture; Natural Resources; Architecture; Cultural Studies; Communications; English; General Studies; Mathematics; Social Sciences; Visual & Performing Arts
Q4:	08, 10, 11, 14, 15, 20, 40, 48, 52	Marketing; Communications Tech.; Engineering; Engineering Tech.; Vocational Home Ec.; Physical Sciences; Preceision Production; Business

Notes: This table presents fields of study at four-year universities grouped into quartiles of major-level exposure. Majors are defined as two-digit Classification of Instructional Programs codes, using NCES crosswalks from 1990 to 2000 and 2010 to make consistent codes. I define major-level shock exposure as the employment-weighted average tariff gap of industries employing recent graduates from that major in the pre-period. Estimations of the effects of exposure to local shocks on enrollment across major quartiles are presented in Table 1.7.

Table 1.29: Foote and Grosz (2020) Community College Major Categories

Category	CIP Code		# of Codes
Information Tech.	10	Communications Technologies/Technicians and Support Services	15
Construction	15	Engineering Technologies/Technicians	54
Manufacturing	46	Construction Trades	22
	47	Mechanic and Repair Technologies/Technicians	34
	48	Precision Production	16
	49	Transportation and Materials Moving	16
Public Services	43	Security and Protective Services	16
	44	Public Administration and Social Service Professions	5
Health	51	Health Professions and Related Clinical Services	196
Business	52	Business, Management, Marketing, and Related Support Services	84
Family/Personal	19	Family and Consumer Sciences/Human Sciences	32
	12	Personal and Culinary Services	25
Education	13	Education	89

Notes: This table is adopted from Foote and Grosz (2020) and presents broad major categories at two-year colleges as defined by groups of two-digit NCES CIP codes. The final column displays the number of individual six-digit CIP code major classifications within the broader two-digit classification. Estimations of the effects of exposure to local shocks on two-year enrollment by major group are presented in Table 1.17.

1.9.3 Policy and Setting Specifics

1.9.3.1 Texas K-12 Finance Policies

Since 1993, Texas has employed a school finance system known as "the Robin Hood Plan." As is often the case with school finance reforms, implementation of this system was the result of a lengthy legal battle. In 1984, Edgewood Independent School District (ISD) and 67 other Texas school districts sued the State over the disparity in resources across the state's school districts. The Texas Supreme Court ruled in favor with the plaintiff districts in *Edgewood v. Kirby* and insisted that conditional on similar levels of local tax effort, districts should receive similar funding levels. Attempts by the state legislature to reform the existing finance system in 1989 and 1991 were struck down for insufficiently addressing disparities and creating an unconstitutional *de facto* state property tax, respectively, prior to the ultimate passage of the Robin Hood plan in 1993. Formally known as the Foundation School Program, the Robin Hood school finance system consists of three tiers of funding, along with a controversial recapture provision.

Tier 1 (Basic Allotment) funding guarantees a minimum level of per-pupil funding to each K-12 district ($2,300 in 1993), conditional on districts meeting a minimum property tax rate threshold ($0.86 per $1,000 in 1993). Districts with insufficient property wealth to meet the minimum funding level at the specified tax rate receive additional state assistance to make up the gap. District basic allotments are distributed per weighted average daily attendance (WADA), a student count measure that adjusts for district and student characteristics representing additional costs.[57]

Tier 2 (Guaranteed Yield) guarantees districts a specified per-WADA funding level ($20.55 in 1993) per penny in property tax effort between the minimum tax rate specified in Tier 1 and a maximum rate ($1.50 in 1993).[58] Property-poor districts that do not meet the guaranteed yield from their own collections receive supplemental state funds to close the gap.

Tier 3 (Facilities) consists of guaranteed-yield funding for facility investments, but unlike Tier 2, operates as a "sum certain" competitive grant program and is not guaranteed to all districts.

Finally, the recapture provision earns the Robin Hood system its name. The state stipulates that districts with per-pupil property wealth that exceeds a specified threshold ($280,000 in 1993) cannot fund their schools with more than the amount equal to their tax rate scaled by the property wealth threshold. The excess property tax revenue is recaptured by the state and redistributed to property-poor districts via Tier 1 and 2.

Hoxby and Kuziemko (2004) discusses efficiency concerns with the Robin Hood system's design relative

[57]District characteristics weighted as additional costs are the average starting salary of teachers in neighboring school districts, the economically disadvantaged student population share, district average daily attendance, location in a rural county, sparsity, and classification as an independent town or small or mid-sized district. Student characteristics weighted as additional costs are participation in special education, compensatory education, career and technology courses, English-Language Learners programs, and gifted and talented programs.

[58]The parameters specifically apply to maintenance and operations tax rates, and from 1999 onward, districts could not use Tier 2 funds to service debt or make facility investments.

to a state property tax that could deliver the same degree of redistribution in K-12 funding. The authors note that the system's reliance on *marginal* gains in property values relative to the per-pupil wealth cutoff exacerbate negative capitalization of the tax in property-rich districts and contend that this decline in property values is unlikely to be fully offset by increases in property-poor districts if these districts were already operating at the efficient level of local K-12 funding. Declining property values in districts above the wealth cutoff would force the state to lower the cutoff in order to raise additional revenue, triggering another round of capitalization. Empirical analyses support these claims: the authors find that although Robin Hood decreased the per-pupil spending gap between property-poor and property-rich districts by $500, it also caused the loss of $27,000 per pupil in property wealth. Efficient confiscation and investment of this amount of wealth by the state could have instead funded all Texas schools at the realized spending level of the top 5% of districts.

1.9.3.2 Generalizability

Table 1.30: K-12 And College 1999 Characteristics: Texas vs. U.S.

	(1) Texas	(2) United States
K-12		
% White	0.431	0.620
% Hispanic	0.396	0.156
% Black	0.144	0.172
Per-pupil Expenditure	$8,165	$7,822
Four-Year Graduation Rate	71%	69%
Two-Year Colleges		
% White	0.522	0.658
% Hispanic	0.291	0.077
% Black	0.121	0.108
Appropriations Per FTE	4,437	4,526
Three-Year Graduation Rate	18%	33%
Four-Year Colleges		
% White	0.620	0.696
% Hispanic	0.195	0.063
% Black	0.089	0.101
Appropriations Per FTE	5,183	5,145
Six-Year Graduation Rate	53%	71%

Notes: This table presents average characteristics for K-12 and college students and schools in Texas and the United States using data from the National Center for Education Statistics Common Core of Data and Integrated Postsecondary Education Data System. All data are from the 1998-1999 school year.

Table 1.31: The Robin Hood Formula Uniquely Shielded TX Students from Local Shocks

	(1) Per-Pupil Property Tax Rev.	(2) Per-pupil Current Exp.	(3) Pass-Through
Texas			
Shock Exposure	-2,786***	39	-0.01
	(421)	(387)	
Percent Change	-75.7	0.43	
Pre-period Mean	3,681	9,094	
N	4,810	4,810	
Rest of U.S.			
Shock Exposure	-322***	-271***	0.84
	(56)	(80)	
Percent Change	-10.2	-2.9	
Pre-period Mean	3,152	9,465	
N	66,045	66,045	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of exposure to negative local labor demand shocks on per-pupil property tax revenues and school spending in TX and the rest of the U.S., using data from the NCES Common Core of Data. Estimates reflect coefficients from a parametric difference-in-differences regression that compares changes in the differences between county-level outcomes in more- and less-exposed counties after the start of the shock (2000) relative to existing trends in these differences. The specification controls for county and year fixed effects, 1990 county characteristics interacted with post, shale and natural gas presence interacted with post, and other changes to U.S. trade policy. Pass-through estimates and column three represent the average decline in school spending per dollar in property tax revenue. Standard errors are clustered by county. Percent changes are calculated as the estimated difference-in-differences coefficient divided by the pre-period mean for high-exposure counties.

CHAPTER 2

Turning around Schools (and Neighborhoods?): School Improvement Grants and Gentrification

This chapter is adapted from "Turning around Schools (and Neighborhoods?): School Improvement Grants and Gentrification" published in *Economics of Education Review* and was reproduced with the permission of the publisher and my co-author Cameron Friday (Friday and Smith, 2023).

2.1 Introduction

K-12 education funding programs often target schools, rather than students directly. The federal government has distributed funding to high-poverty schools since 1965 through the Title I grant program, totalling $15 billion per year over the past decade. Because schools in the United States are broadly segregated by race and income (McGrew, 2019), programs that aim to increase resources available to disadvantaged students allocate additional funding to schools with large minority or low-income student populations. However, households may respond to shocks to school spending when choosing schools and neighborhoods, potentially reducing the effectiveness of such targeting. In particular, sorting in response to investments in local schools can result in housing price increases that reflect households' valuation of the investments (Tiebout, 1956; Bayer et al., 2007; Cellini et al., 2010; Bayer et al., 2020). Household sorting of this nature may result in less effective targeting of education funding if disadvantaged families are priced out of neighborhoods zoned for newly improved schools.

We examine whether household sorting in response to changes in school funding inhibits spending from reaching targeted students with a case study of the effect of federal grants to low-performing schools in Metro-Nashville Public Schools (MNPS). We study the School Improvement Grant program (SIG), a Title I program that invested over $7 billion in the nation's lowest-achieving schools between 2009 and 2016. Using real estate assessment and property transaction records linked to MNPS attendance zones, we find that home values in SIG-receiving school zones increased by 10.5% more than that of houses just outside the attendance zone boundary in response to the investment. Under the assumption that no other factors affecting house prices changed across attendance zone boundaries differentially before and after SIG treatment, our results indicate that households were willing to pay $3 for every dollar in per-pupil grant aid to live in SIG school zones.

This price increase is consistent with funding for education in SIG schools previously being suboptimally low (Oates, 1969) but also represents a potential barrier for low-income families to access newly improved schools. We characterize sorting in response to SIG and explore its consequences using Consumer Financial Protection Bureau (CFPB) mortgage data, eviction counts from the Eviction Lab at Princeton University, and school enrollment demographics. We show that following program roll-out homebuyers moving into previously majority-minority SIG attendance zones reported 9% higher income and were 13% more likely to be white. At the same time, the share of students at SIG schools who are white increased by almost 20%. In tandem with evidence of a 35% increase in evictions in neighborhoods zoned for treated schools and a 15% decline in nonwhite enrollment at SIG schools, our results suggest that SIG-induced gentrification displaced disadvantaged residents from their previous neighborhoods and schools.

We identify effects of school improvement grants on housing prices using a difference-in-differences design that appeals to the boundary-discontinuity estimations used in previous literature (Black, 1999; Schwartz et al., 2014). Exploiting variation in access to schools implementing SIG-funded interventions across space and time, we compare changes in prices of homes sold within a half-mile of attendance zone boundaries for SIG-treated schools. Causal interpretation rests on the assumption that home values of properties on either side of attendance zone boundaries for SIG schools would evolve in parallel in the absence of the program, which is supported by common trends in sale prices prior to SIG funding receipt. Our preferred specification restricts the sample to single-family home and duplex sales and includes boundary segment-by-year fixed effects that account for time-varying neighborhood amenities. However, our results are robust to a variety of alternative specifications. We provide evidence that our estimations are not confounded by the Great Recession and housing crash, rising housing prices common to all neighborhoods zoned for low-performing schools, and changes in neighborhood demographics, but note with caution that we cannot rule out bias from other unobservable neighborhood amenities that differentially change just inside of SIG school zones compared to just outside these zones.

Next, we examine how sorting in response to SIG funding affected the demographic composition of neighborhoods and schools. To do so, we match CFPB mortgage data and eviction counts from the Eviction Lab to MNPS attendance zones. Difference-in-differences estimates that compare Census tracts or block groups within SIG attendance zones to those outside these zones before and after SIG receipt indicate that the reported income of mortgage applicants in neighborhoods zoned for SIG schools increased by 9% within two years of grant receipt and that the share of white home buyers rose by 13%. Enrollment data from the National Center for Education Statistics (NCES) reveal similar patterns of racial integration into SIG schools, with the share of students who are white increasing by almost 20% relative to that at non-SIG schools. These results suggest that SIG funding led to moderate integration by income and race in previously majority-minority neighborhoods. However, this influx of higher-income households into SIG neighborhoods also displaced some disadvantaged residents: evictions increased by 35% in neighborhoods zoned for SIG schools relative to non-SIG attendance zones after grant receipt, and nonwhite enrollment at SIG schools declined by 15%.

SIG likely had heterogeneous effects on housing prices across housing markets, and our primary estimates represent local average treatment effects from SIG's implementation in Nashville, a booming housing market. We provide support for the generalizability of our findings by examining the household sorting response to SIG in California, a state where we observe eligibility and funding receipt for all SIG cohorts and which successfully used SIG funding to improve student achievement in low-performing schools (Friday, 2021). Matching data on housing prices, mortgage characteristics, and evictions to school attendance zones in California, we estimate a difference-in-differences specification that compares changes in neighborhoods zoned for schools that received SIG funding to those zoned for SIG-eligible schools that did not receive grants. We find evidence of willingness-to-pay for school funding and neighborhood sorting that is broadly consistent with the results from our case study: for every $1,000 in per-pupil SIG funding, home values in California increased by 3% and the the reported income of homebuyers in neighborhoods zoned for SIG schools increased by 3.5%.

Our analysis advances existing research on the relationship between school characteristics and home values by estimating the capitalization of SIG funding (and any associated changes in school quality) into local housing prices. Prior research finds that housing prices increase by 3-10% in response to a standard deviation increase in school achievement[1] and found that households are willing to pay more than $1 for every dollar increase in K-12 spending (Barrow and Rouse, 2004; Cellini et al., 2010; Bayer et al., 2020). We contribute to this literature by estimating willingness-to-pay for school spending targeting schools at the bottom of the achievement distribution, where additional funding may be especially impactful. Our estimates of over a $3 increase in house price for a dollar in per-pupil grant aid suggest that education funding among this subgroup falls well below the efficient level of provision (Oates, 1969).

Our results also add to a growing literature examining how changes in the provision of K-12 education influence neighborhood sorting. Notable policies addressing the inequitable provision of K-12 education include desegregation, school choice, and funding increases targeting low-performing schools. Baum-Snow and Lutz (2011) show that court-ordered school desegregation efforts induced white-flight into suburban school districts. Moreover, previous research shows that both "exit options" (Schwartz et al., 2014; Zheng, 2019) and "forced choice" (Wigger, 2020) varieties of school choice weaken the link between neighborhood school characteristics and home values.[2] If our results are interpreted as causal, we extend this literature by showing that investments to improve low-performing schools can make neighborhoods served by those schools more desirable and attract wealthier households. Together, these findings indicate the need to consider the potential for neighborhood sorting and displacement when designing K-12 finance and assignment policies aiming to make schooling more equitable.

More broadly, this paper contributes to literatures on the relationship between public good provision and housing markets and the nature of gentrification of low-income neighborhoods. Following SIG-funded interventions, wealthier and whiter residents moved into neighborhoods zoned for SIG schools. Integration into SIG school zones is consistent with the pattern of sorting found in response to other education-related policies (Billings et al., 2017) and broader place-based policies targeting low-income neighborhoods (Diamond and McQuade, 2019). We contribute to these findings by testing whether whether sorting displaces existing residents of neighborhoods benefiting from increased public good provision. In contrast to previous literature

[1] See, for instance, Black (1999), Kane et al. (2006), Bayer et al. (2007), Black and Machin (2011), Machin (2011), Dhar and Ross (2012), Gibbons et al. (2013), Schwartz et al. (2014), Collins and Kaplan (2017), and Caetano (2019).

[2] Wigger (2020) defines "exit options" as choice policies that provide families alternatives to their assigned neighborhood schools and "forced choice" as assignment policies that require all families to submit school choice applications and do not utilize neighborhood school boundaries.

on gentrification, which found little empirical evidence of displacement in gentrifying neighborhoods (McKinnish et al., 2010; Disalvo, 2022), our results suggest that gentrification displaced a non-negligible share of existing residents of SIG neighborhoods via increased evictions. Taken together, our findings illustrate a major limitation of place-based public good provision: sorting, or concurrent gentrification, may displace the initially targeted population.

2.2 The School Improvement Grant Program

2.2.1 Program Overview

In an effort to boost achievement in the nation's "persistently lowest-achieving schools," the American Recovery and Reinvestment Act of 2009 allotted $3 billion to the School Improvement Grants (SIG) program to fund aggressive school turnaround programs. An additional $4 billion in funding for SIG between 2010 and 2016 followed, and by 2012 the SIG program had invested up to $6 million per school in more than 1,300 of the country's lowest-achieving schools (Department of Education, 2012).

The federal government first allotted SIG funds to state education agencies (SEAs) via grants based on existing Title I funding formulas. To receive funding, states submitted applications to the Department of Education (Ed) with identified SIG-eligible schools, complying with Ed's three tiers of eligibility. Tier I schools, the highest priority for SIG funds, were comprised of the lowest-achieving five percent of Title I schools in improvement, corrective action, or restructuring in the state. Similarly, Tier II schools consisted of the lowest-achieving five percent of schools eligible for, but not receiving, Title I funds for school improvement. All remaining Title I schools in improvement, corrective action, or restructuring were designated as Tier III schools. SEAs distributed 95% of SIG funds to local education agencies (LEAs; i.e., school districts) to implement school turnaround programs in eligible schools, prioritizing funding toward districts with Tier I and II schools.

Schools awarded SIGs were to use the grant money to implement one of four schoolwide intervention models, each with the dual purpose of disrupting the status quo (i.e., making substantial changes to school operations and staff) and increasing school resources (Zimmer et al., 2017). 95% of SIG schools implemented either the "transformation model" (75%) or the "turnaround model" (20%) (Ginsburg and Smith, 2018).[3] Both interventions required replacing the principal, implementing a teacher evaluation system accounting for student achievement growth, and increasing learning time; the turnaround model further mandated the replacement of at least 50% of school staff. Increasing learning time intuitively should lead to learning gains, and student achievement also likely benefits from the replacement of low-performing teachers if new teachers are more effective, which becomes more likely with achievement-based evaluations. Empirical evidence supports this notion: in a meta-analysis of school turnaround interventions, Schueler et al. (2020) identify extended learning times and teacher replacement as characteristics of interventions associated with greater

[3]The other 5% of SIG schools implemented either the "restart model," which handed over schools to a charter management organization, or the "closure model," which shut down low-performing schools and allowed previously assigned students to enroll in higher-achieving schools within the district.

effects on student achievement.

Research on school improvement grants generally finds immediate positive effects of the program on student achievement in various local and statewide settings (Friday, 2021; Sun et al., 2020; Carlson and Lavertu, 2018; Sun et al., 2017), and studies with extended time horizons show that test score gains last beyond the three-year intervention (Sun et al., 2020) and may even increase over time (Friday, 2021). Identified underlying mechanisms include SIG treatment reducing unexcused absences, improving retention of effective teachers, and developing greater teacher professional capacity (Sun et al., 2017). A notable exception to the literature's consensus, nationwide analysis in Dragoset et al. (2017) finds no significant effect of SIG-funded interventions on math or reading test scores, high school graduation, or college enrollment. However, the study only analyzes achievement data from 2012 and 2013 and does not estimate longer-term effects.

2.2.2 SIG and Household Sorting

Through increasing both perceived and real school quality, SIG-funded interventions may disrupt existing housing market equilibria in settings where school assignment depends on household location. We outline how this may take shape, appealing to the model of public good provision developed by Tiebout (1956).

Canonical theory posits that households choose communities to reside in based in part on the provision of local public goods, "voting with their feet" to reveal their preferences (Tiebout, 1956). Researchers have often applied this framework to study the relationship between observable measures of school quality and housing prices,[4] taking advantage of the U.S.'s general reliance on neighborhood schools that admit only (or at least guarantee spots for) students residing in a specific attendance zone. Households that value school quality sort into neighborhoods zoned for high-quality schools, yielding an equilibrium where, all else equal, housing prices across school attendance zones reflect differences in the marginal household's willingness to pay for school quality. A robust literature provides evidence for the capitalization of school quality into housing prices in the U.S., generally finding that a one-standard deviation increase in test scores raises home values by 2-4% percent (Black, 1999; Black and Machin, 2011; Machin, 2011; Gibbons et al., 2013), although a price gradient as large as 10% for a standard deviation increase in test scores has been found in urban settings (Kane et al., 2006).

Households may observe SIG-funded interventions and their effects on school quality through multiple channels. In our setting of Davidson County, local newspapers reported on the receipt of SIG funding (Barnes, 2015; Gonzales, 2015a), as well as specific aspects of school reform prescribed by SIG interventions such as principal replacement (Gonzales, 2015b,c) and extending learning time (Beecher, 2014). If parents expect SIG-funded interventions will improve low-performing schools, or more generally value school inputs, then housing prices should respond fairly quickly to the announcement of grant receipt and local news coverage. However, parents may instead update beliefs over school quality when they observe changes in observable measures such as standardized test scores (Figlio and Lucas, 2004). Organizations such as GreatSchools

[4] E.g., Black (1999), Kane et al. (2006), Bayer et al. (2007), Black and Machin (2011), Cellini et al. (2010), Machin (2011), Dhar and Ross (2012), Gibbons et al. (2013), Schwartz et al. (2014), Collins and Kaplan (2017), Caetano (2019), and Bayer et al. (2020).

publish simple school ratings based on test score levels and growth to inform parents choosing schools. GreatSchools partners with Zillow (and did so at the time of SIG implementation in MNPS), so that households can easily observe school characteristics when searching for homes. Existing literature suggests that the household sorting response to SIG should increase over time. In particular, Sun et al. (2017) find an increase in the popularity of SIG schools on school choice applications in San Francisco that grows over time and mirrors the pattern of achievement effects found by the authors. This supports the notion that SIG interventions are salient enough to elicit behavioral responses from households when they improve student outcomes.

Sorting in response to SIG may inhibit funding from reaching targeted students and induce broad change of neighborhood demographics. In particular, relatively wealthy households with substantial willingness to pay for access to SIG-funded schools may price out low-income households of neighborhoods zoned for these newly improved schools. Billings et al. (2017) find sorting of relatively high-income households into neighborhoods zoned for previously low-performing schools in Charlotte, NC following school closures under No Child Left Behind that gave school choice priority to students residing within attendance zones of newly closed schools. On a larger scale, Bayer et al. (2020) exploit exogenous variation in school spending from court-mandated school finance reforms across the country and estimate that school poverty rates decreased by 0.21% in response to a 1% increase in school spending. SIG may elicit a similar response by drastically increasing school spending over three years to implement a turnaround reform.

2.2.3 SIG and Metro-Nashville Public Schools

This paper examines the neighborhood sorting response to SIG funding receipt in Nashville, Tennessee. Serving Nashville and the surrounding area that constitutes Davidson County, Metro Nashville Public Schools (MNPS) consisted of 67 elementary schools, 28 middle schools, and 12 high schools as of the 2009-2010 school year, when SIG was expanded by the ARRA. Students are zoned to attend neighborhood schools based on their residence and follow "pathways" across school levels in MNPS, such that all students from a given elementary school feed into the same middle school, and all students from said middle school move onto the same high school. MNPS offers multiple forms of "exit option" school choice, including magnet schools and charter schools, which potentially allow for families dissatisfied with their assigned neighborhood school to move to a different school. However, at the beginning of the twenty-first century the predominant form of choice was to leave the district altogether: between 2000 and 2012, MNPS lost nearly 10 percent of its student population annually to private schools and other public school districts. Simultaneously, the percent of the MNPS student population qualifying for free or reduced-price meals rose from 45 percent to over 70 percent.

When the SIG program was expanded under the Obama Administration, MNPS presented a clear need to address its lowest-performing schools. Achievement across the district was poor: the percentage of students in grades 3-8 meeting proficiency in Math was just 28%. Furthermore, the district's low-performers fell even further below district goals. Proficiency rates on standardized tests were below 20% in reading and 10% in math at seven middle schools.

MNPS received school improvement grants in the 2012-2013 and 2015-2016 school years to implement intervention models in two cohorts of low-performing schools. The 2012 cohort consisted of three middle schools (Brick Church, Gra-Mar, and John Early Paidea Magnet) and three elementary schools (Buena Vista, Napier, and Robert Churchwell Museum Magnet).[5] Two middle schools and one elementary school implemented the turnaround model, while the remaining two middle schools and elementary school carried out the transformation model. The 2015 cohort consisted of two elementary schools expanding early learning programs (Inglewood and John Whitsitt), two middle schools implementing the turnaround model (Jere Baxter and Madison), a high school implementing the transformation model (Pearl Cohn), and an elementary school undergoing a restart as a charter school (KIPP at Kirkpatrick).[6] Figure 2.1 depicts geographic variation in SIG treatment across the district, and Figure 2.8 portrays the location of traditional and charter schools in MNPS.

[5]Notably, two treated schools are magnet schools. Since these do not map to attendance zones, we do not examine them in our housing price analysis. The presence of desirable magnet schools that do not require residence in specific neighborhoods for enrollment would attenuate estimates of SIG on housing prices by offering households zoned for untreated schools a degree of access to SIG schools. Still, magnet school admission in MNPS is not guaranteed, in contrast to zoned neighborhood schools.

[6]Although not initially one of the four prescribed turnaround programs, SIG grants in later cohorts could be used to add or expand pre-k and kindergarten programs at elementary schools through the "Early Learning" module.

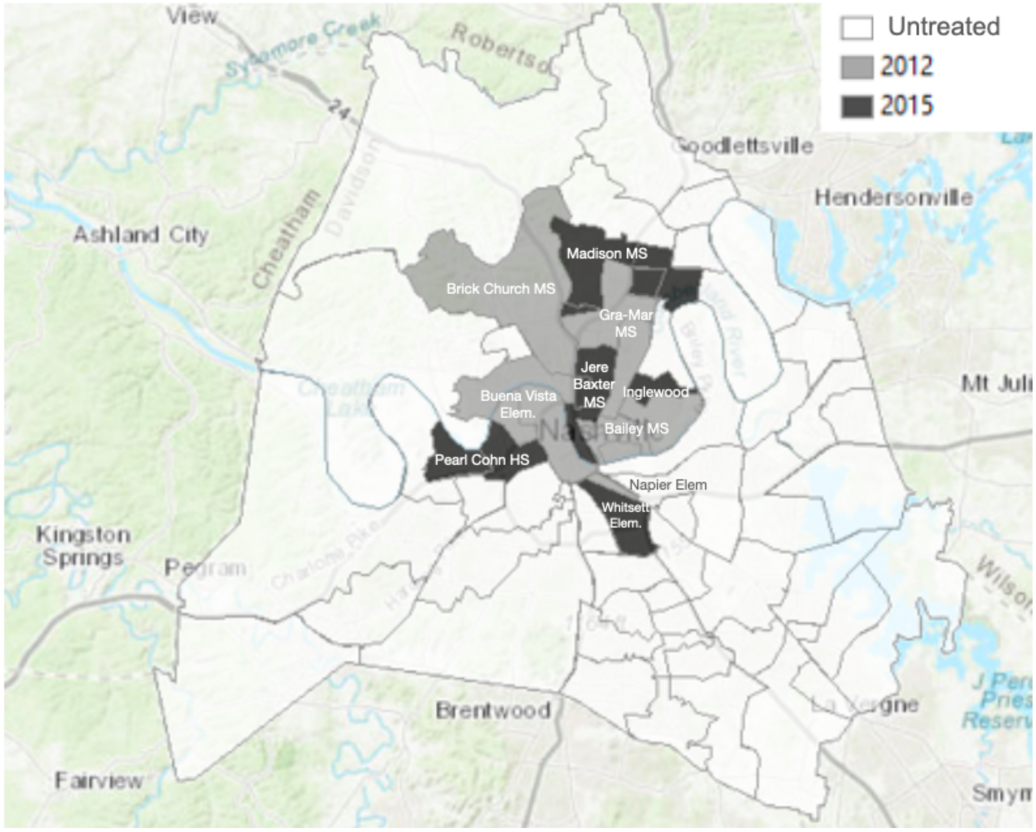


Figure 2.1: MNPS School Zones: SIG Intervention Year

Notes: This figure maps MNPS school zones, denoting attendance zones for schools receiving school improvement grants (labeled by recipient school) in the 2012-2013 and 2015-2016 school years. Each attendance zone corresponds to a pathway of assigned elementary, middle, and high schools.

Because the Tennessee Department of Education (TDOE) switched testing regimes in 2015, rendering comparisons of student achievement using publicly available school-level testing data before and after the change uninformative, we are limited in our ability to estimate the effects of SIG-funded interventions on test scores in MNPS.[7] We instead rely on the previously mentioned literature as evidence of the "first-stage" of our analysis. We also speak to one potential mechanism through which SIG improved outcomes. By year two of interventions, the 2012 cohort of schools in MNPS had increased their school year length by an average of 47.8 hours, equivalent to an additional week of instruction and a 3.9% increase relative to the 2011-2012 mean.

[7] Appendix Table 2.15 presents difference-in-differences estimates of the test score effects of SIG for the 2012 cohort in MNPS, using school-by-grade achievement data from TDOE. Although noisy, estimates suggest that SIG interventions improved the lower tail of achievement at treated schools, as evident by reductions the share of students scoring "Below Basic" in reading and math.

2.3 Data and Descriptive Statistics

2.3.1 Davidson County Property Sales

We create a panel of Davidson County parcels and sales by year using property assessment and sales data obtained from the Metro Nashville Davidson County Division of Assessments. We observe the sales price, transaction date, and parcel location for all parcel transfers from 2000 until 2019. Through linking sales records to the most recent prior county assessment, we observe parcel and housing characteristics at the time of sale (e.g. the number of stories and bedrooms). We drop all sales with missing sales prices or those of less than $1,000, limiting our sample to "arms-length" transactions. Furthermore, we restrict our analysis to single-family homes and duplexes, the most likely forms of housing utilized by households with school-age children. Our final panel includes 86,651 sales involving 41,116 unique properties.

In order to designate SIG treatment to homes, we first match parcels with a geocoded map of school assignment zones for Metro Nashville Public Schools. We use zones from a year prior to the initial cohort of SIG in MNPS as our baseline, avoiding bias from any potentially endogenous changes in zoning following SIG. After geocoding our sales data through the Census Bureau's geocoding service, we match each parcel to their assigned elementary, middle, and high school. Furthermore, we calculate the distance between parcels and assignment zone boundaries, allowing us to limit our sample to homes close to boundaries of SIG-treated schools during analysis.

Homes are assigned treatment based on whether their zoned elementary, middle, or high school implement a SIG-funded intervention. Table 2.1 presents summary statistics and balance tests for homes sold from 2000 to 2011 (all sample years prior to the implementation of SIG interventions for the earliest cohort) zoned for schools that later received SIG treatment compared to those zoned for untreated schools. Columns 1, 2, and 3 correspond to sales across our entire sample in 2010, while Columns 4, 5, and 6 limit sales to those of parcels within a half-mile of SIG school zone boundaries. Homes zoned for SIG schools sold for over $55,000 less (in 2010 dollars) than those zoned for non-SIG schools prior to treatment. This difference can partially be explained by the smaller size of homes in SIG school zones, as measured by the number of rooms, baths, finished square footage, and property acreage. A joint F-test confirms what is clear from the differences across individual factors: homes sold in the pre-period significantly differed across observables by their treatment status.

Table 2.1: Davidson County Sales Data

	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample			Close to Boundary		
	Untreated	SIG	Diff	Untreated	SIG	Diff
Sale Price	180,987	125,216	-55,771***	138,963	115,400	-23,562***
	(139,282)	(101,423)	(1,084)	(85,441)	(73,624)	(1,358)
ln(Sale Price)	11.886	11.489	-0.397***	11.691	11.473	-0.219***
	(0.659)	(0.727)	(0.006)	(0.563)	(0.642)	(0.010)
Bedrooms	3.034	3.005	-0.029***	2.969	2.928	-0.041***
	(0.794)	(0.841)	(0.007)	(0.768)	(0.809)	(0.013)
2+ Stories	0.112	0.106	-0.006**	0.071	0.085	0.014***
	(0.315)	(0.307)	(0.003)	(0.257)	(0.279)	(0.005)
Rooms	6.454	6.163	-0.290***	6.135	5.960	-0.175***
	(1.747)	(1.620)	(0.015)	(1.509)	(1.488)	(0.025)
Baths	1.778	1.724	-0.055***	1.658	1.672	0.014
	(0.815)	(0.711)	(0.007)	(0.694)	(0.692)	(0.012)
Half Baths	0.272	0.199	-0.072***	0.187	0.161	-0.026***
	(0.467)	(0.420)	(0.004)	(0.408)	(0.384)	(0.007)
Finished Sq. Feet	1,775	1,569	-206***	1,562	1,501	-61***
	(854)	(653)	(7)	(638)	(643)	(11)
Parcel Acreage	0.442	0.301	-0.140***	0.395	0.335	-0.061***
	(0.931)	(0.511)	(0.007)	(1.044)	(0.585)	(0.014)
Building Age	41.123	45.072	3.950***	41.286	41.371	0.085
	(26.528)	(31.229)	(0.258)	(25.959)	(28.181)	(0.460)
Observations	25,793	25,827	51,620	7,262	6,808	14,070

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents summary statistics and balance tests for homes sold from 2000 - 2011 zoned for schools that later received SIG treatment compared to those zoned for untreated schools. Columns 1 through 3 correspond to the entire sample of 2010 sales in Davidson County, while Columns 4 through 6 limit sales to parcels within a half-mile of attendance zone boundaries for SIG schools. P-values for Columns 3 and 6 represent standard t-tests of mean equality. Sales prices are in 2010 dollars, and the underlying data were obtained from the Metro Nashville Davidson County Division of Assessments.

2.3.2 Neighborhood Composition

We further characterize the household sorting response to SIG using data on neighborhood composition from multiple sources. Baseline neighborhood demographics are observed in the 2010 Decennial Census, which we link to treatment data by matching geocoded maps of Census block groups and MNPS school zones. Table 2.2 depicts average household characteristics for block groups zoned for untreated and SIG schools. SIG and non-SIG neighborhoods consisted of starkly contrasting compositions of races: non-SIG school zones were 76.3% white, while SIG zones were 56.1% Black. Households in SIG neighborhoods typically housed more occupants than those in non-SIG zones, and untreated and SIG neighborhoods also differed in the type of housing utilized: households in SIG block groups were 13.4 percentage points more likely to be renter-occupied. Median household income in non-SIG school zones nearly doubled that of SIG zones, consistent with other stark socioeconomic differences. Finally, the last two rows of Table 2.2 present summary statistics of data from the Eviction Lab at Princeton University, which has made aggregations of individual records of eviction cases from courts across the country publicly available. SIG block groups had approximately triple the number of court-ordered evictions and eviction filings per 1,000 residents as non-SIG block groups.[8] Notably, differences in observable neighborhood characteristics substantially shrink when we limit the sample to block groups contiguous to SIG attendance zone boundaries. For example, the difference in median household income falls from $27,812 to $4,456.

[8] An eviction filing may result in a court ruling for or against eviction or a settlement between the landlord and tenant.

Table 2.2: Davidson County 2010 Occupied Housing Demographics

	(1)	(2)	(3)	(4)	(5)	(6)
	All Block Groups			Contiguous Block Groups		
	Untreated	SIG	Diff	Untreated	SIG	Diff
% White	0.763	0.367	-0.395***	0.694	0.452	-0.242***
	(0.190)	(0.249)	(0.010)	(0.203)	(0.242)	(0.019)
% African American	0.142	0.561	0.419***	0.183	0.484	0.301***
	(0.159)	(0.274)	(0.009)	(0.188)	(0.265)	(0.019)
% Households with 3+ Occupants	0.323	0.357	0.033***	0.279	0.320	0.041***
	(0.139)	(0.109)	(0.007)	(0.149)	(0.087)	(0.013)
% Renter Occupied	39.076	52.515	13.440***	53.396	49.802	-3.595**
	(26.889)	(20.827)	(0.920)	(25.803)	(22.899)	(1.685)
Median Household Income	59168	31356	-27812***	37391	32935	-4456***
	(38113)	(11636)	(1,223)	(17037)	(12525)	(1,066)
Median Gross Rent	817.852	676.294	-141.557***	724.408	654.334	-70.073***
	(363.861)	(230.955)	(12.289)	(222.556)	(255.341)	(16.298)
Eviction Rate	6.062	17.102	11.039***	8.899	18.845	9.946***
	(9.770)	(15.626)	(0.634)	(10.188)	(20.590)	(1.571)
Eviction Filings Rate	11.768	34.325	22.557***	20.059	38.006	17.947***
	(20.114)	(36.094)	(1.375)	(27.452)	(46.797)	(3.783)
Observations	4,149	1,207	5,356	718	416	1,134

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents summary statistics and balance tests for demographics from Davidson County blocks in the 2010 Decennial Census. Columns (1) through (3) reflect the entire county, while columns (4) through (6) limits the sample to "contiguous" block groups, defined as those that are at least partially contained within a half-mile buffer around SIG attendance zone boundaries. Block groups that do not completely fall on one side or the other of the boundary (i.e., those that are partially treated) are tossed out. P-values for Columns 3 and 6 represent standard t-tests of mean equality.

To examine sorting in response to SIG, we use loan-level Consumer Financial Protection Bureau mortgage data, available from 2007 to 2017. Although the exact address of homes are suppressed to maintain privacy, individual mortgages are identified by Census tracts, which we match to school zones. The data include both characteristics of the applicant (race, gender, income) and the loan (amount, agency, property type, approval status). By observing applicant income and race, we can characterize households moving into SIG school zones following grant receipt and test if the sorting response is consistent with gentrification. Table 2.3 presents summary statistics and balance tests for mortgages of homes sold in SIG and non-SIG school zones from 2007 to 2011, demonstrating that homebuyers in SIG neighborhoods received smaller loans, reported lower income, and were more likely to be Black prior to treatment.

Table 2.3: Davidson County Mortgage Characteristics, 2007-2011

	(1) Untreated	(2) SIG	(3) Diff
Loan Amount	180,816	121,510	-59,306***
	(36,715)	(30,624)	(6,685)
Applicant Income	75,234	51,344	-23,890***
	(16,752)	(17,078)	(3,630)
White Share of Borrowers	0.721	0.545	-0.176***
	(0.163)	(0.298)	(0.061)
Black Share of Borrwowers	0.125	0.311	0.186***
	(0.167)	(0.315)	(0.064)
Observations	25	186	532

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents summary statistics and balance tests for pre-SIG mortgage characteristics using data from the Consumer Finance Protection Bureau. Individual mortgages are identifiable at the census tract-level, and census tracts are coded as treated (SIG) if they are fully contained within school attendance zones for SIG schools; similarly, groups are coded as untreated if they are fully contained within attendance zones of non-SIG schools. P-values for Column 3 represent standard t-tests of mean equality.

2.3.3 School Characteristics

We utilize school-level data on student achievement and demographics from the Tennessee Department of Education and National Center for Education Statistics Common Core of Data to present baseline differences in characteristics of SIG and non-SIG schools. Along with enrollment demographics, we observe the share of students in grades 3-8 that score in each of four achievement categories of the state's standardized exams in math and reading, Below Basic, Basic, Proficient, and Advanced, since the 2009-2010 school year. As shown in Table 2.4, student demographics starkly differed between SIG and untreated schools in 2010. SIG elementary schools served an almost exclusively Black population, while 44.9% of students at non-SIG elementary schools were Black–a slim plurality. The student populations of SIG and untreated middle schools similarly differed along demographics (79.4% and 37.6% Black, respectively). Student achievement across MNPS was quite low; still, SIG schools stood out as the lowest performers. Two-thirds of students at SIG middle schools scored Below Basic in math, compared to 47% of students in non-SIG middle schools.

Table 2.4: MNPS 2010 Demographics and Achievement

	(1)	(2)	(3)	(4)	(5)	(6)
	Elementary Schools			Middle Schools		
	Untreated	SIG	Diff	Untreated	SIG	Diff
Total Enrollment	183	164	-19	568	430	-138
	(80)	(42)	(57)	(144)	(91)	(105)
% White	0.350	0.016	-0.333**	0.393	0.129	-0.264**
	(0.224)	(0.009)	(0.160)	(0.145)	(0.035)	(0.105)
% African American	0.449	0.962	0.513***	0.376	0.794	0.417***
	(0.267)	(0.010)	(0.190)	(0.120)	(0.094)	(0.088)
% Free or Reduced Lunch	0.746	0.948	0.201	0.712	0.923	0.210
	(0.219)	(0.042)	(0.156)	(0.168)	(0.005)	(0.122)
% Below Basic RLA	0.180	0.477	0.297***	0.221	0.349	0.128**
	(0.081)	(0.149)	(0.060)	(0.075)	(0.038)	(0.055)
% Below Basic Math	0.174	0.422	0.249***	0.470	0.664	0.193*
	(0.084)	(0.036)	(0.060)	(0.132)	(0.071)	(0.096)
% Basic RLA	0.473	0.427	-0.045	0.443	0.487	0.044
	(0.093)	(0.066)	(0.066)	(0.064)	(0.028)	(0.046)
% Basic Math	0.465	0.462	-0.003	0.339	0.275	-0.064*
	(0.089)	(0.010)	(0.063)	(0.048)	(0.068)	(0.036)
% Proficient RLA	0.266	0.087	-0.179***	0.287	0.144	-0.143**
	(0.087)	(0.081)	(0.062)	(0.082)	(0.001)	(0.060)
% Proficient Math	0.259	0.101	-0.157***	0.135	0.046	-0.089**
	(0.077)	(0.017)	(0.055)	(0.058)	(0.004)	(0.042)
% Advanced RLA	0.082	0.009	-0.073	0.049	0.020	-0.029
	(0.077)	(0.002)	(0.055)	(0.048)	(0.012)	(0.034)
% Advanced Math	0.103	0.014	-0.089	0.055	0.015	-0.041
	(0.076)	(0.009)	(0.054)	(0.045)	(0.006)	(0.032)
Observations	61	2	63	19	2	21

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents summary statistics and balance tests for demographics and standardized test scores at SIG and non-SIG schools in MNPS. Columns 1 through 3 correspond to elementary schools, while Columns 4 through 6 correspond to middle schools. There were no treated high schools in the initial SIG cohort at MNPS. P-values for Columns 3 and 6 represent standard t-tests of mean equality. The underlying data were obtained from the National Center for Education Statistics Common Core of Data and the Tennessee Department of Education.

2.4 Methods

We exploit variation in access to SIG-funded schools across space and time via a boundary-discontinuity difference-in-differences design to identify the capitalization of school improvement grants into housing prices. To motivate this design, consider a canonical hedonic model of housing markets:

$$ln(P_{iz}) = \beta Q_z + \gamma X_{iz} + \varepsilon_{iz} \tag{2.1}$$

where the price of house i in attendance zone z P_{iz} is a function of school quality Q_z and observable house characteristics X_{iz}. If unobserved house or neighborhood characteristics are correlated with school quality, then estimates of β will be biased. Researchers have addressed this endogeneity issue by extending the regression discontinuity design to analysis of housing markets, noting that school quality discontinuously changes at school attendance zone boundaries (Black, 1999; Gibbons et al., 2013; Schwartz et al., 2014). The "boundary discontinuity design" is often specified through boundary segment fixed effects ω_k, along with restricting the sample to homes close (e.g., within a half-mile) to shared attendance zone boundaries.

To isolate the effects of school improvement grants from preexisting differences in school quality, we follow Schwartz et al. (2014) in extending the boundary-discontinuity approach to a dynamic setting. One way to do so would be to estimate the following two-way fixed effects (TWFE) specification:

$$ln(P_{izkt}) = \alpha_0 + \gamma X_{izkt} + \alpha_z + \alpha_t + \omega_k + \beta^{TWFE} SIG_{izt} + \varepsilon_{izkt} \tag{2.2}$$

where α_z accounts for time-invariant differences in home prices across school zones, α_t captures time-varying shocks to prices that affect the entire sample–such as the 2007 housing crash, and SIG_{izt} is an indicator variable that equals one for sales of homes zoned for SIG schools occurring after grant receipt.

Goodman-Bacon (2021c) shows that when treatment timing varies, estimates of β^{TWFE} will be biased if treatment effects are dynamic.[9] There are multiple reasons to believe SIG-funded interventions will generate dynamic treatment effects in housing markets. Grant funding through the program is received over three years to implement turnaround interventions. If families value school inputs, they may desire access to treated schools immediately upon funding receipt. On the other hand, parents may not base beliefs in school quality off of school inputs, but rather school output, in which case demand for housing in SIG-zoned schools would only increase if and when observable achievement metrics improve. Previous literature suggests that improvements in achievement from SIG interventions increase over time (Friday, 2021; Sun et al., 2017), and as the observable improvements in achievement become more salient, households may increasingly desire access to SIG schools. If this holds true, comparisons of homes zoned for the 2015 cohort of SIG schools to those of the 2012 cohort would bias estimates of β^{TWFE}.

[9] In particular, Goodman-Bacon (2021c) shows that when treatment timing varies, estimates of β^{TWFE} partially reflect comparisons of newly treated units to previously treated units. If treatment effects are dynamic, this use of previously treated units as a control group typically biases estimates of β^{TWFE} away from the sign of the true treatment effect. Intuitively, units still *responding to treatment* are not a valid counterfactual to represent potential outcomes in the *absence of treatment*.

To bypass the bias identified by Goodman-Bacon (2021c), we follow Deshpande and Li (2019) and Flynn and Smith (2021) in estimating a stacked difference-in-differences specification to identify effects of SIG-funded interventions on housing prices, an estimation strategy that will make no comparisons based off of variation in treatment timing. To implement this, we first re-organize our data into two "stacks," representing our two cohorts of treated schools. The first stack is comprised of homes sold in attendance zones for schools receiving SIG grants in 2012, labeled as treated homes, and homes sold in attendance zones for schools that never receive a SIG grant, labeled as untreated homes. Similarly, the second stack is comprised of homes sold in attendance zones for schools receiving SIG grants in 2015, labeled as treated homes, and homes sold in attendance zones for schools that never receive a SIG grant, labeled as untreated homes. We further restrict the 2012 stack to consist of only homes close to attendance zone boundaries for schools treated in 2012 and similarly restrict the 2015 stack to consist of only homes close to attendance zone boundaries for schools treated in 2015, following the boundary discontinuity literature (Black, 1999; Gibbons et al., 2013; Schwartz et al., 2014).[10]

Figure 2.2 visually portrays the identifying variation, mapping MNPS school zone boundaries with half-mile buffers surrounding attendance zone boundaries that separate zones for schools receiving school improvement grants in the 2012-2013 and 2015-2016 school years from untreated schools. The first stack makes comparisons of homes within a half-mile of boundaries for schools treated in 2012 (dark boundary segments, light buffer), and the second stack makes comparisons of homes within a half-mile of boundaries for schools treated in 2015 (light boundary segment, dark buffer)[11]. Our working data set for estimations consists of these two stacks appended together.

[10]We also omit boundaries along rivers or highways, following Black (1999) in excluding borders that physically divide neighborhoods.

[11]We refer to school years by their fall calendar year (i.e., calling the 2012-2013 school year "2012").

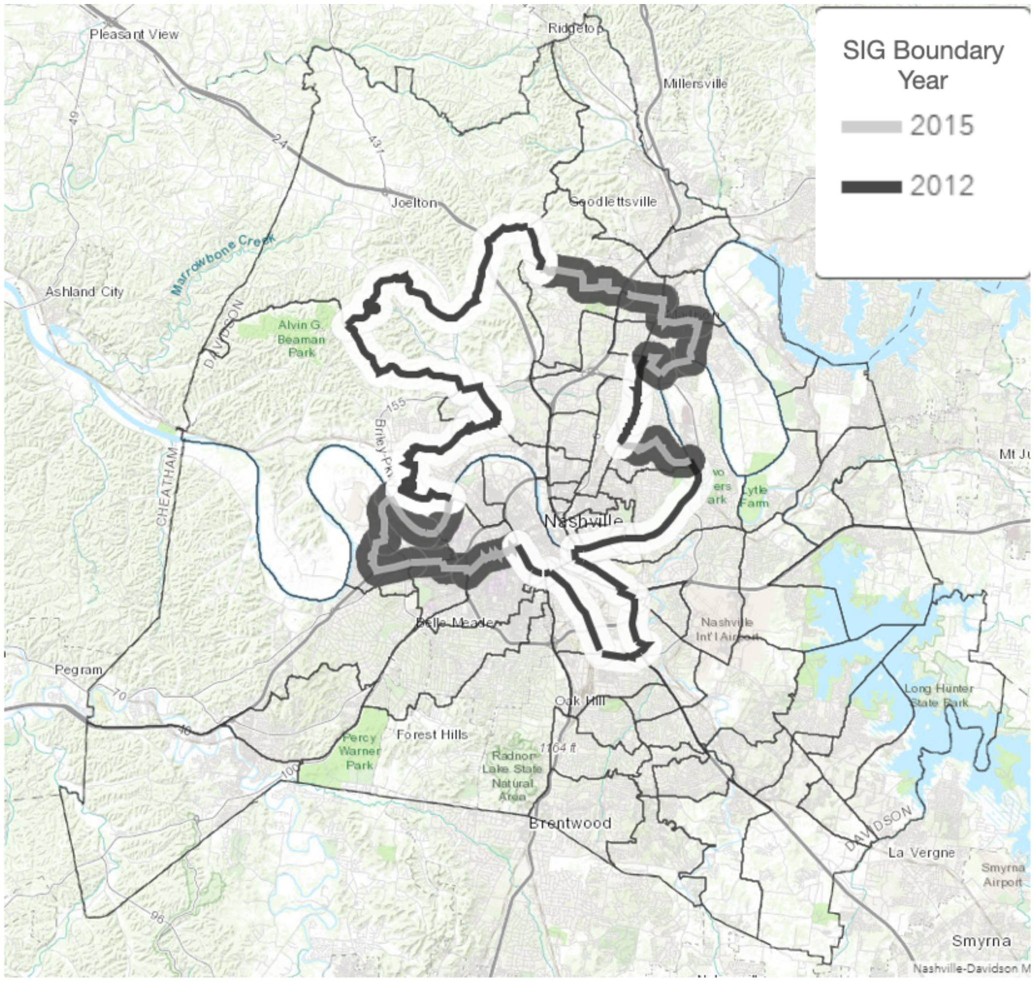


Figure 2.2: Stacked DD Geographic Variation

Notes: This figure maps MNPS school zone boundaries, tracing out half-mile buffers surrounding attendance zone boundaries that separate schools receiving school improvement grants in the 2012-2013 and 2015-2016 school years from untreated schools. Our stacked difference-in-differences estimation compares changes in sale prices of homes located just inside attendance zone boundaries for schools receiving SIG grants in to untreated schools located just outside these boundaries. The first stack makes comparisons of homes within a half-mile of boundaries for schools treated in 2012 (dark boundary segments, light buffer) and the makes comparisons of homes within a half-mile of boundaries for schools treated in 2015 (light boundary segment, dark buffer).

The stacked difference-in-differences estimation separately estimates a standard difference-in-differences for each cohort of SIG schools, which compares changes in the sale price of homes located just inside SIG attendance zones to untreated homes located just outside SIG attendance zones, and averages the two simple difference-in-differences coefficients from the 2012 and 2015 stack. Formally, we estimate the following specification:

$$ln(P_{iczkt}) = \alpha_0 + \gamma X_{izkt} + \alpha_{cz} + \alpha_{ct} + \omega_{ck} + \beta^{stacked} SIG_{iczt} + \varepsilon_{iczkt} \tag{2.3}$$

where i, c, z, k, and t index homes, stacks, school attendance zones, boundary segments, and years, respectively. Because sale price P_{cizlt} is determined in part by the physical qualities of a unit, we control for X_{izkt}, a vector of home characteristics observed in the most recent assessment that includes the number of bedrooms and bathrooms, the age of the home, and the size of both the lot and home itself. All fixed effects, α_{cz}, α_{ct}, and ω_{ck} are indexed by both stack and another characteristic (school zone, year, and border segment, respectively). Segmenting fixed effects by stack ensures that $\beta^{stacked}$ is only identified off of within-stack (i.e., within cohort) variation. Specifically, this prevents $\beta^{stacked}$ from reflecting any comparison that uses already-treated homes as comparisons.[12] With this in mind, α_{cz} and α_{ct} serve the same purpose as their TWFE counterparts, school zone and year fixed effects: to capture time-invariant differences in home prices across school zones and time-varying shocks to prices that affect all of Nashville, respectively. Moreover, stack-by-boundary-segment fixed effects ω_{ck} account for time-invariant neighborhood amenities that plausibly benefit houses on either side of an attendance zone boundary, and augmenting the specification to instead include stack-by-boundary-segment-by-year fixed effects ω_{ckt} will account for changes over time in neighborhood amenities common across an attendance zone boundary. Applying the intuition of Black (1999), we specify boundary fixed effects as dummy variables indicating to which segment of the attendance zone boundaries of stack c's treated schools a particular home is located closest. Finally, SIG_{cizt} is an indicator variable that equals one for sales of homes zoned for treated schools in stack c occurring after the intervention begins.

Interpreting β as the causal effect of a SIG-funded intervention on home values requires the identifying assumption that in the absence of the program, house prices would evolve in parallel on either side of attendance zone boundaries for would-be treated schools. Although this assumption is fundamentally untestable, an implication is that prices evolve similarly leading up to the start of SIG funding receipt. To test this, and to flexibly estimate the full dynamics of SIG treatment, we specify the following stacked event study:

$$ln(P_{iczkt}) = \alpha_0 + \gamma X_{izkt} + \alpha_{cz} + \alpha_{ct} + \omega_{ck} + \sum_{\substack{y=-11 \\ y\neq -1}}^{7} \pi_y 1\{t - t_c^* = y\} * treat_{ciz} + \varepsilon_{izkt} \tag{2.4}$$

where t_c^* is the initial year of SIG funding receipt in stack c (2012 for c = 1 and 2015 for c = 2) and $treat_{ciz}$ is a dummy variable indicating if a parcel is zoned for a school treated in its stack. Estimates of π_y represent the difference in prices of homes on either side of attendance zone boundaries for treated schools, relative to this difference one year prior to grant receipt. Examination of trends in home prices along SIG boundaries in the pre-period may be particularly important in our setting given our treatment timing relative to the housing

[12] Although Goodman-Bacon (2021c) identifies comparisons of newly to previously treated units as a potential source of bias, yet-to-be treated units can serve a valid counterfactual for newly treated units (under the assumption of exogenous treatment timing). In our setting, however, exploiting this type of comparison would come at a cost. Since the later-treated cohort began interventions only three years after the earlier cohort, we would have to restrict our post period to three years to maintain a balanced panel. If housing prices are sticky, this may not be a long enough post period to detect price responses to SIG-funded interventions.

crash and onset of the Great Recession.[13]

The appeal of using boundary discontinuity designs to identify the capitalization of school characteristics into housing prices lies in the notion that homes located within small distances of school attendance boundaries (on either side) likely benefit from similar unobservable neighborhood amenities. Although we cannot test if unobservables vary across attendance zone boundaries for treated schools, we can provide support for the notion that untreated homes located just outside of SIG school zones are valid counterfactuals for those located just inside treated zones by testing if observable characteristics discontinuously change at SIG boundaries. Table 2.5 presents coefficients from separately regressing observable characteristics of homes sold between 2012 and 2019 (after the first cohort of MNPS was implemented) on an indicator variable for whether or not a home was zoned for a SIG-treated school, year fixed effects, and attendance zone boundary segment fixed effects. Although some coefficients are significant at conventional levels, their magnitudes are small relative to sample means. Furthermore, column (8) shows that a hedonic housing price index that summarizes overall housing quality in a single measure does not economically or statistically significantly differ across SIG attendance boundaries.[14]

Table 2.5: Housing Characteristics Along SIG Boundaries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Acreage	Finished SF	Building Age	Two+ Stories	Rooms	Bedrooms	Bathrooms	ln(HPI)
.5 mi bandwidth								
SIG	-0.0915***	-41.77	-8.771*	0.105*	-0.252*	-0.00851	0.0362	0.00794
	(0.0317)	(37.88)	(4.303)	(0.0573)	(0.142)	(0.0607)	(0.0514)	(0.0327)
N	14368	14368	14368	14368	14368	14368	14368	14368

Standard errors clustered by boundary segment (23 clusters) are in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table depicts if and how characteristics of homes on either side of SIG attendance boundaries differed differed. Estimates reflect coefficients from separately regressing observable characteristics (acreage, finished square feet, age of the building, a dummy for if the house had at least two stories, the number of rooms, the number of bedrooms, the number of bathrooms, and a hedonic housing price index that summarizes overall housing quality in a single measure.) on an indicator variable for whether or not a home was zoned for a SIG-treated school, year fixed effects, and attendance zone boundary segment fixed effects. We limit the sample to single-family homes and duplexes sold between 2012 and 2019 (after the first cohort of MNPS was implemented) within a half-mile of a SIG attendance zone border.

[13]Mean reversion from the housing crash could result in substantial increases in housing prices in our pre-period or, if sufficiently delayed, in our post-period; this would be of particular concern if the bite of the crash was stronger in SIG school neighborhoods than in control areas. Although the variation in treatment timing in our research design should guard against mean reversion from the 2007 housing crash and recovery from overpowering housing price estimations, as a falsification exercise we formally test whether home values on either side of SIG boundaries were more affected by the crash with an event study that estimates the evolution of the difference in housing prices just inside to just outside SIG boundaries relative to this difference at the start of the housing crash. Figure 2.9provides evidence that housing prices evolved similarly on both sides of SIG boundaries through the crash and Great Recession. Another potential concern is that the Great Recession may have more harshly affected school and household resources in SIG schools and neighborhoods, resulting in test score declines that pushed schools into SIG eligibility. We show similar trends in district-level expenditures in MNPS, Shelby County Schools (Memphis), and the rest of the state in Figure 2.10 through the Great Recession. Although we cannot directly rule out the recession affecting household resources differentially in SIG neighborhoods, we note that such a dynamic would primarily be of concern for our displacement analysis. We find no differential changes in eviction rates in SIG and non-SIG school zones through the Great Recession in Figure 2.11, suggesting that the recession does not explain our displacement results in Section 2.5.2.

[14]The housing price index is constructed from estimating a linear regression of sale price on the other characteristics presented in Table 2.5 with pre-period data for all sales in Davidson County.

To help contextualize magnitudes of our main house price estimations, we estimate the capitalization of school quality prior to SIG implementation in MNPS (2000 to 2011). Table 2.6 presents coefficients from boundary discontinuity specifications with the natural log of sale price as the dependent variable. Columns (1) and (2) show the pre-treatment difference in housing prices along SIG attendance zone borders, indicating that homes just inside attendance zones that later implemented a SIG-funded intervention sold for 16.6 log points less than homes just on the other side of attendance zone boundaries. Columns (3) and (4) instead follow previous literature in using elementary school test scores as a proxy for school quality. We follow Black (1999) in constructing a standardized index based off the sum of school-level achievement (proficiency rates in our setting) on math and reading standardized tests. Estimates shows that a one standard deviation increase in elementary school achievement was associated with a 10.4% higher sale price, a nearly identical magnitude to that found by Kane et al. (2006) in the comparable setting of Charlotte, North Carolina.

Table 2.6: Pre-Treatment School Quality Capitalization

	(1) Pre-SIG	(2) + Controls	(3) Test Scores	(4) + Controls
SIG	-0.197**	-0.166**		
	(0.0717)	(0.0657)		
Proficiency Rates			0.139**	0.0993**
			(0.0562)	(0.0373)
N	14070	14070	13622	13622

Standard errors clustered by boundary segment (23 clusters) are in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table depicts how sale prices of single-family homes and duplexes differed along SIG school zone borders prior to treatment (2000-2011). Estimates reflect coefficients from separately regressing observable characteristics the natural log of sale price on an indicator for SIG treatment (Columns 1 and 2) and on a standardized measure of elementary school test scores (Columns 3 and 4). All estimations control for year and boundary fixed effects, and columns 2 and 4 control for observable home characteristics. Coefficients for Columns 3 and 4 correspond to the change in housing price from a one standard deviation increase in school-level elementary test scores, and the sample size is smaller for test score estimations because of missing data for two schools. Estimations in Columns 3 and 4 also include high school fixed effects, so that differences in secondary school quality to not bias estimates of capitalization of elementary school quality.

2.5 Results

2.5.1 Capitalization of School Improvement Grants

We begin by presenting event study coefficients from estimating equation 2.4 in Figure 2.3, limiting our sample to single family homes and duplexes within a half-mile of attendance zone boundaries separating SIG and non-SIG school zones. Home values evolve in parallel prior to the start of SIG interventions, suggesting that they would continue to do so in the absence of the program. Following treatment, prices of homes zoned for SIG schools immediately increase relative to their neighbors across the boundary. Furthermore, this increase grows over time, matching the dynamics of the achievement effects found by previous research on SIG (Friday, 2021; Sun et al., 2017) and the dynamics of housing price effects found in literature analyzing other K-12 policies (Bayer et al., 2020; Wigger, 2020). These treatment effect dynamics confirm that two-way fixed estimates would be biased and support the use of a stacked design. Moreover, an immediate response that grows over time suggests that parents respond both to the receipt of SIG funding (changes in school inputs) and presumptive improvements in test scores (changes in school outputs).[15]

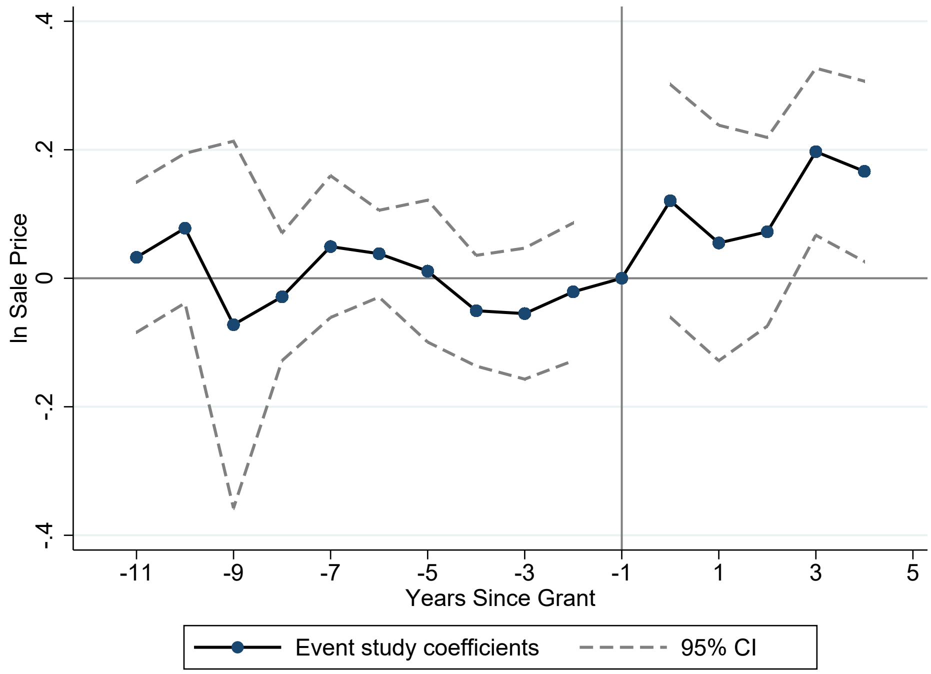


Figure 2.3: Housing Price Stacked Event Study

Notes: This figure plots π_y coefficients from equation 2.4, a stacked event study specification that compares changes in the natural log of sale price (in 2010 \$) of homes located on either side of attendance zones of schools implementing SIG-funded interventions. The sample is limited to single family homes and duplexes located within a half-mile of attendance zone boundaries separating treated and untreated school zones. Dashed lines represent the 95% confidence interval of estimates based off of robust standard errors clustered by boundary segment-by-post.

We summarize the dynamic effects seen in Figure 2.3 by estimating equation 2.3, a stacked difference-in-differences (DD) specification. Table 2.7 contains estimates of $\beta^{stacked}$ from our preferred specification and a variety of robustness checks. Inference is based off of robust standard errors clustered at the boundary segment-by-post level, appealing to the notion that our identifying variation comes from the discontinuous access to SIG treatment across attendance zone boundaries in the post-period but not the pre-period. Panel A reflects estimates on a sample of single-family homes and duplexes, while Panel B restricts the sample only to single-family homes. Our base model estimates a stacked DD on a sample of sales of homes located

[15]Although due to data limitations we are unable to confirm that SIG-funded interventions improved test scores in Nashville, existing literature finds positive positive effects of SIG on student achievement in in various local and statewide settings (Friday, 2021; Sun et al., 2020; Carlson and Lavertu, 2018; Sun et al., 2017) and positive effects of other school turnaround programs on student achievement in Tennessee, including in MNPS (Pham et al., 2020).

within a half-mile of SIG attendance boundaries, comparing changes in sale prices of homes zoned for SIG-treated schools to those just on the other side of the attendance zone boundary. Columns (1) and (2) estimate this specification with and without controlling for observable home characteristics, yielding coefficients in Panel A that indicate SIG-funded interventions increased housing prices by 10.2 (p-value .144) and 10.0 (p-value of .112) log points, respectively. To account for time-varying neighborhood amenities common to both sides of attendance zone boundaries, we include boundary segment-by-year fixed effects in Columns (3) onward. Doing so greatly improves the precision of our estimates, and our estimate of 10.0 log points in Column 3 is statistically different from zero at the 1% level of significance. That our estimate changes little when controlling for observable home characteristics suggests that homes sold on either side of SIG attendance borders had similar characteristics. Still, differential changes in the composition of unobservable characteristics of homes sold across SIG attendance boundaries could bias estimates. To address this concern, we include parcel fixed effects in Column (4), identifying the estimate of β from repeat sales of the same homes and finding a similar coefficient of 10.2 log points (p-value of .011). Because we expect SIG treatment to increase neighborhood desirability by improving the quality of targeted schools, increases in home values should be driven by households with school-age children. As a proxy for the presence of children, we limit our sample to only homes with at least three bedrooms in Column (5). Finally, Columns (6) and (7) narrow our bandwidth to limit the sample to homes sold within .35 and .25 miles of SIG school zone boundaries, respectively.

Table 2.7: Capitalization Stacked DD: SIG-funded Interventions Raise House Prices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	.5 mi	+ Controls	+ Boundary X Year FEs	+ Parcel FEs	3+ BR	.35 Mile	.25 Mile
A. SF & Duplexes							
SIG	0.102	0.0999	0.100^{***}	0.102^{**}	0.0822^{**}	0.0723^{***}	0.0413^{*}
	(0.0687)	(0.0618)	(0.0305)	(0.0381)	(0.0352)	(0.0232)	(0.0219)
N	22134	22134	22083	17505	16285	15109	10255
B. Single-Family Only							
SIG	0.112	0.106^{*}	0.108^{***}	0.118^{***}	0.0848^{**}	0.0814^{***}	0.0586^{***}
	(0.0690)	(0.0622)	(0.0311)	(0.0360)	(0.0354)	(0.0241)	(0.0210)
N	20648	20648	20599	16400	14956	14163	9544

Robust standard errors clustered by boundary segment-by-post (46 clusters) are in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of β from equation 2.3, a stacked difference-in-difference specification that compares changes in the logged sale price of homes sold just on either side of SIG attendance zone boundaries. All estimations include school zone-by-stack and stack-by-year fixed effects. Panel A includes both single-family homes and duplexes, and Panel B shows robustness to restricting only to single-family homes. Column (1) limits the sample to homes within a half-mile of a SIG attendance zone boundary and includes boundary segment fixed effects. Column (2) adds controls for observable home characteristics. Columns (3) onward use boundary segment-by-year fixed effects to account for time-varying neighborhood amenities common across both sides of a school zone boundary. Column (4) includes parcel fixed effects, so that the estimate of β is identified by repeat sales. Column (5) limits the sample to homes with at least three bedrooms. Finally, Columns (6) and (7) restrict the sampling bandwidth to .35 and .25 miles from SIG attendance zone boundaries, respectively.

To further characterize the price effects shown in Table 2.7, we explore whether SIG treatment of elementary, middle, or high schools had heterogeneous effects on housing prices. Table 2.8 presents coefficients from separately estimating our stacked difference-in-differences on samples limited to homes within a half-mile of attendance zone boundaries for SIG-treated elementary (Columns 1 and 2), middle (Columns 3 and 4), and high schools (Columns 5 and 6). Consistent with prior literature that finds stronger capitalization of elementary and high school quality (Caetano, 2019), our estimates suggest that SIG-funded interventions in elementary or high schools resulted in larger housing price increases than those in middle schools. We interpret the high school result with caution, however, because our sample only features one treated high school.

Table 2.8: Capitalization Stacked DD: Heterogeneity By School Level

	(1)	(2)	(3)	(4)	(5)	(6)
	Elem.	+ Controls	MS	+ Controls	HS	+ Controls
SIG	0.195^{***}	0.181^{***}	0.0719	0.0722	0.228^{**}	0.236^{**}
	(0.0504)	(0.0440)	(0.0453)	(0.0454)	(0.0784)	(0.0768)
N	5183	5183	13872	13872	5374	5374

Robust standard errors clustered by boundary segment-by-post are in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of β from estimations of equation 2.3, a stacked difference-in-difference specification that compares changes in sale price of homes sold just on either side of SIG attendance zone boundaries, that separately limit the sample to boundary segments for SIG-treated elementary, middle, and high school attendance zones, respectively. All estimations include school zone-by-stack and stack-by-boundary-segment-by-year fixed effects. All coefficients reflect a sample of single-family homes and duplexes sold within .5 miles of a SIG attendance zone boundary. Columns (2), (4), and (6) include controls of observable home characteristics.

Our setting allows for a natural application of Oates's (1969) empirical test for whether a local public good is efficiently provided. Appealing to the Tiebout (1956) model of local public good provision based on households 'voting with their feet,' Oates hypothesized that an increase in home values in response to a marginal increase in locally financed public goods would indicate that households value the marginal \$1 in expenditure more than the marginal \$1 lost via taxation and that the local public good had been underprovided. Because SIG was financed completely by the federal government and required no increase in local taxation, in our setting the Oates test translates to testing whether home values increase by more or less than \$1 for every \$1 in increased expenditure.[16] Based on the mean pre-treatment sale price of a SIG-zoned home of \$125,216, estimates from our preferred specification (Column 3) imply an increase in home value of \$13,148 in response to a SIG intervention. Compared to the average per-pupil grant of over \$3,818, the greater than \$3 willingness-to-pay for every additional \$1 in per-pupil funding is consistent with funding

[16]We implicitly assume that SIG funding leads to a dollar-for-dollar increase in school spending (i.e., no crowd-out); Friday (2021) finds evidence of this at the district level.

for SIG-receiving schools previously being suboptimally low.[17] We note with caution that this interpretation requires the assumption that no other local amenities are changing differentially along SIG school zone boundaries.

The above interpretation requires the assumption that no other local amenities are changing differentially along SIG school zone boundaries at the same time as treatment. Although this assumption is fundamentally untestable, we examine testable implications of violations dye to potential confounders–namely, concurrent gentrification. We first conduct three sets of falsification exercises to support a causal claim of our estimates. Because Nashville sustained a booming housing market over the past decade, one may be concerned that all low-income neighborhoods experienced large price increases, so that estimates from Table 2.7 do not reflect the capitalization of improved school quality from SIG interventions.[18] We note two sets of attendance zone boundaries that should not experience a changing price gradient if our observed housing price increases are caused by SIG: attendance zone boundaries separating untreated elementary schools *within the attendance zone of a SIG-treated middle or high school* and boundaries for schools that were *eligible* for SIG but *did not receive grants*. Both of these sets of attendance zone boundaries represent actual changes in school asssignment, along with potentially other neighborhood characteristics, but homes on both sides experience the same SIG treatment status. Appendix Figure 2.12 imposes attendance zone boundary segments used for these two falsification exercises over a map of MNPS attendance zones. For a third set of falsification exercises, we construct placebo SIG attendance zone boundaries by shifting the real boundaries 1, 3, and 5 miles in each cardinal direction. These placebo boundaries do not represent changes in school assignment or other neighborhood characteristics except for those that arise by chance.

Table 2.9 presents estimates from the first two falsification exercises. Coefficients in panel A reflect estimates from a stacked difference-in-differences specification that compares changes in sale price of homes sold just on either side of attendance zone boundaries for elementary schools within the same SIG-treated middle school or high school. We assign placebo "treatment" to the elementary school within each treated upper school with the lowest 2010 standardized test scores, mimicking the true SIG treatment assignment mechanism. Coefficients in Panel B reflect estimates of a difference-in-differences specification with placebo "treatment" defined as an indicator equalling unity for sales occurring in 2012 or after for homes zoned for elementary schools that were eligible for SIG but did not receive grants to implement turnaround programs. We find null results across all variants of these two exercises. Table 2.17 presents estimates from the final set of falsification estimations. Only one out of twelve difference-in-differences estimates statistically differs from zero, consistent with the frequency we would expect to find a statistically significant result purely from chance. Also consistent with randomness, the estimates vary in sign and magnitude both within and across directions and distances.

[17]The willingness-to-pay estimates associated with each specification from Table2.7 all yield this same conclusion from an Oates (1969) efficiency test.

[18]In particular, Guerrieri et al. (2013) predict that during to a housing boom, high-income households expand their housing consumption by migrating to poorer areas adjacent to high-income neighborhoods.

Table 2.9: Falsification Exercises: House Prices Do Not Change Where They Should Not

	(1)	(2)	(3)	(4)	(5)	(6)
	.5 mi	+ Controls	+ Boundary X Year FEs	3+ BR	.35 Mile	.25 Mile
A. Within-Zone Boundaries						
Placebo SIG	0.00311	0.00326	0.00403	0.00352	0.00345	0.00267
	(0.00610)	(0.00625)	(0.00659)	(0.00767)	(0.00735)	(0.00647)
N	15537	15537	15523	11786	11378	7892
B. SIG Eligibility						
Placebo SIG	-0.0259	-0.0259	-0.0443	-0.0699	-0.0630	-0.162
	(0.0791)	(0.0791)	(0.0575)	(0.0595)	(0.0633)	(0.103)
N	16815	16815	16777	11555	11686	7959

Robust standard errors clustered by boundary segment-by-post are in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of two sets of falsification exercises. Coefficients in Panel A reflect estimates of β from a stacked difference-in-difference specification that compares changes in sale price of homes sold just on either side of attendance zone boundaries for elementary schools *within the same SIG-treated middle school or high school zone*. Placebo "treatment" is assigned to the elementary school within each treated upper school with the lowest 2010 standardized test scores, mimicking the true SIG treatment assignment mechanism. Coefficients in Panel B reflect estimates of a difference-in-difference specification with placebo "treatment" defined as an indicator equalling unity for sales occurring in 2012 or after for homes zoned for schools that were eligible for SIG but did not receive grants to implement turnaround programs. All estimations include school zone-by-stack and stack-by-year fixed effects. Column (1) limits the sample to homes within a half-mile of a SIG attendance zone boundary. Column (2) adds controls for observable home characteristics. Columns (3) onward include boundary segment-by-year fixed effects that account for time-varying neighborhood amenities common across attendance zone boundaries. Column (4) limits the sample to homes with at least three bedrooms. Finally, Columns (5) and (6) restrict the sampling bandwidth to .35 and .25 miles from SIG attendance zone boundaries, respectively.

The falsification exercises suggest that no confounding shocks systematically improved housing prices in low-income neighborhoods zoned for low-performing schools throughout Nashville, but they do not eliminate the possibility of confounding amenity changes specifically in SIG neighborhoods. Because we find evidence in Section 2.5.2 that increased demand to live in SIG school zones is accompanied by changing neighborhood demographics–particularly, that white and wealthy homebuyers disproportionately move in, the estimates in Table 2.7 may reflect a combination of valuation of improved school quality and of increasingly desirable neighbors. We present estimates in Table 2.18 from stacked difference-in-differences estimations that additionally control for two time-varying neighborhood-level controls from Consumer Financial Protection Bureau mortgage data: the natural log of the average reported income of homebuyers and the share of homebuyers that are white. Although these endogenous characteristics represent "bad controls," in that they are outcomes that might also be affected by SIG, we are reassured by finding that their inclusion does not affect our capitalization estimates (Angrist and Pischke, 2009). We note with caution that other changes to unobservable neighborhood amenities that differentially change just inside of SIG school zones compared to just outside these zones may still be reflected in changing housing prices.

Finally, we discuss the plausibility of the magnitudes of our housing price estimates within the context of other literature. Our preferred specification indicates that following SIG funding receipt, home values just inside SIG school zone boundaries rose by $13,148 relative to those just across the boundary. Although this magnitude may seem large, back-of-the-envelope calculations based on existing estimates of test-score capitalization and direct estimates of SIG's effect on test scores from the literature support its plausibility under the assumption that SIG had a similarly sized effect on test scores in MNPS as found in other settings. Specifically, our point estimate falls between the seminal test-score capitalization estimates of Black (1999) scaled by SIG's causal effect on test scores found by Sun et al. (2020) ($9,814) and those found by Friday (2021) ($19,886).[19] The salience of SIG-funded comprehensive turnaround interventions, which often required schools to replace principals and 50% of school staff and were covered by local media (Barnes, 2015; Gonzales, 2015b,c,a; Beecher, 2014), likely explains the immediacy of the housing price response. The estimated effect of a 10.5% increase in housing prices also resembles the magnitude of effects of salient interventions affecting school quality in comparable settings of Memphis, TN (Collins and Kaplan, 2017), Denver, CO (Wigger, 2020), and Charlotte, NC (Billings et al., 2017),[20] and SIG joins capital bonds (Cellini et al., 2010) and school finance reforms (Bayer et al., 2020) as sources of K-12 spending increases that yielded greater willingness-to-pay estimates of more than a dollar for a dollar increase in spending.

2.5.2 Effects on Neighborhood and School Composition

In addition to bidding up prices of homes zoned for treated schools, sorting in response to SIG interventions may have altered the composition of SIG-zoned neighborhoods and schools. We examine how characteristics of households moving into treated neighborhoods changed in response to SIG-funded interventions using mortgage data from the Consumer Financial Protection Bureau. Although we cannot observe property addresses, mortgages are identifiable at the Census-tract level, which we match to MNPS school zones. We adapt our stacked difference-in-differences design to fit tract-level mortgage data, estimating the following event study and DD specifications:

$$y_{cat} = \alpha_0 + \alpha_{ca} + \alpha_{ct} + \sum_{\substack{y=-4 \\ y\neq -1}}^{2} \pi_y 1\{t - t_c^* = y\} * treat_{ca} + \varepsilon_{cat} \tag{2.5}$$

$$y_{cat} = \alpha_0 + \alpha_{ca} + \alpha_{ct} + \beta^{stacked} SIG_{cat} + \varepsilon_{cat} \tag{2.6}$$

where c, a, and t index stacks (i.e. cohort of SIG), Census tracts, and years, respectively.[21] Interpreting β as causal requires the identifying assumption that in the absence of SIG treatment, outcomes would evolve in

[19] Black (1999) finds a willingness-to-pay of $9,039 (2010$) for a school-level deviation (approximately 5%) increase in test scores. Sun et al. (2020) finds that SIG led to a 0.228 student-level standard deviation increase in math scores by the intervention's third year. We translate this effect to 1.09 school-level standard deviations using Kane et al.'s (2006) student-level to school-level standard deviation conversion ratio of 0.21:1. Friday (2021) finds that SIG led to an 11% improvement in school proficiency rates. We scale the $9,039 willingness-to-pay by 1.09 and 2.2 (11%/5%) to reach $9,814 and $19,886, respectively.

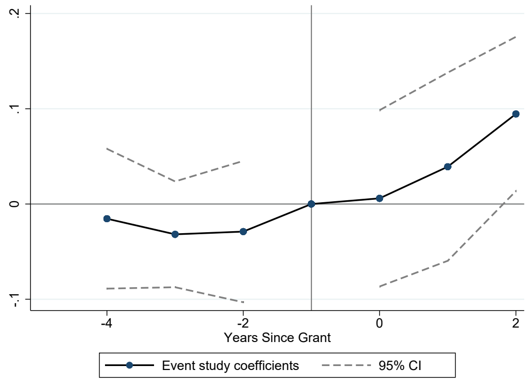
[20] Collins and Kaplan (2017) finds a 7.8% increase in sale price of homes subject to redistricting in Memphis, TN, Wigger (2020) finds that housing values for homes previously assigned to higher rated neighborhood schools fall by up to 9.3% when reassigned to larger "forced choice" zones in Denver, CO, and Billings et al. (2017) finds that housing prices in the "highest-quality" neighborhoods that qualify for priority in enrollment lotteries for oversubscribed schools increase by 8.8%.

[21] Because we only observe mortgage data from 2007 to 2017, to maintain a balanced panel our event time is limited to run from -4 to 2.

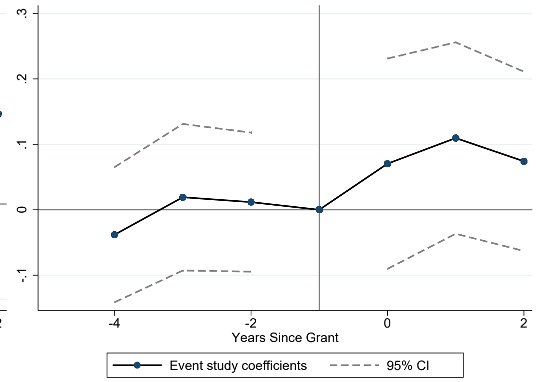
parallel for Census tracts zoned for SIG non-SIG schools.

Figure 2.4 presents event study coefficients for mortgage characteristics of interest: the natural logs of loan amount (a) and applicant income (b) and the shares of white (c) and Black (d) homebuyers. Each panel displays flat pre-trends, suggesting that outcomes in SIG and non-SIG Census tracts would continue to evolve in parallel in the absence of treatment. Panel (a) confirms the primary result of our capitalization analysis. Following the start of SIG-funded interventions, loan amounts increase in Census tracts zoned for SIG schools relative to those zoned for non-SIG schools–a response that grows over time. This phenomenon corresponded with an increase in reported applicant income, and appears to be driven by the differential migration of white households into SIG neighborhoods, as evident by panels (b), (c), and (d). DD estimates of these effects, presented in Table 2.10, show moderately sized effects of SIG on homebuyer characteristics, all of which are statistically significant at at least the 10% level.[22]

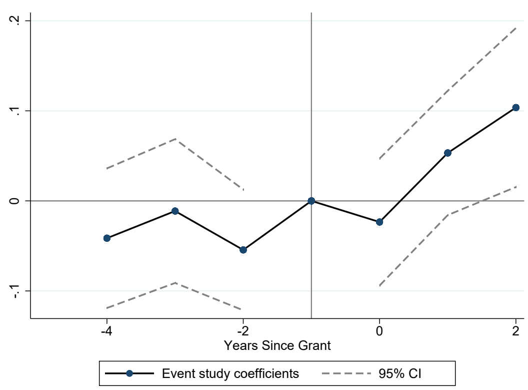
[22]The estimated effect of SIG on loan amount as measured by mortgage data may seem inconsistent with that from our full capitalization analysis (it is noticeably smaller). This is due to only observing mortgage data through 2017–recall that our main estimate of house price effects grow over time. When limiting the post period in our capitalization estimations to also be only two years, we obtain a DD estimate of housing price effects that is statistically indistinguishable from the loan amount estimate in Table 2.10.



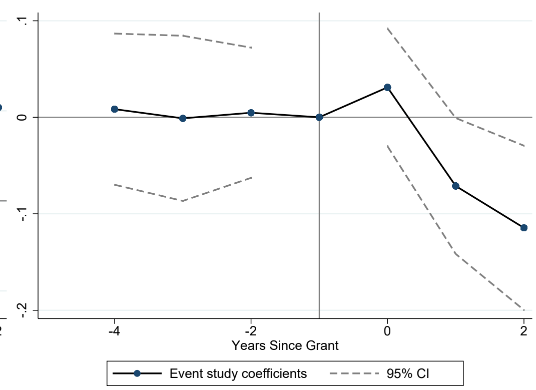
(a) ln(Loan Amount)



(b) ln(Applicant Income)



(c) Share White Buyers



(d) Share Black Buyers

Figure 2.4: Stacked Event Studies: Homebuyer Characteristics

Notes: These figures present estimates of the effect of SIG-funded interventions on neighborhood composition, as measured by Consumer Financial Protection Bureau mortgage data identifiable at the Census tract. Estimates reflect coefficients from a stacked event study regression that compares changes in characteristics of tracts zoned for schools treated as part of the 2012 or 2015 SIG cohorts to those of untreated schools. The specification controls for tract-by-stack and year-by-stack fixed effects.

Table 2.10: Mortgage Characteristic DDs: Higher-Income Whites Move to SIG Neighborhoods

	(1)	(2)	(3)	(4)
	ln Loan Amount	ln Applicant income	% White Buyers	% Black Buyers
SIG	0.0657*	0.0859*	0.0710***	-0.0545***
	(0.0355)	(0.0509)	(0.0231)	(0.0182)
N	1661	1661	1661	1661

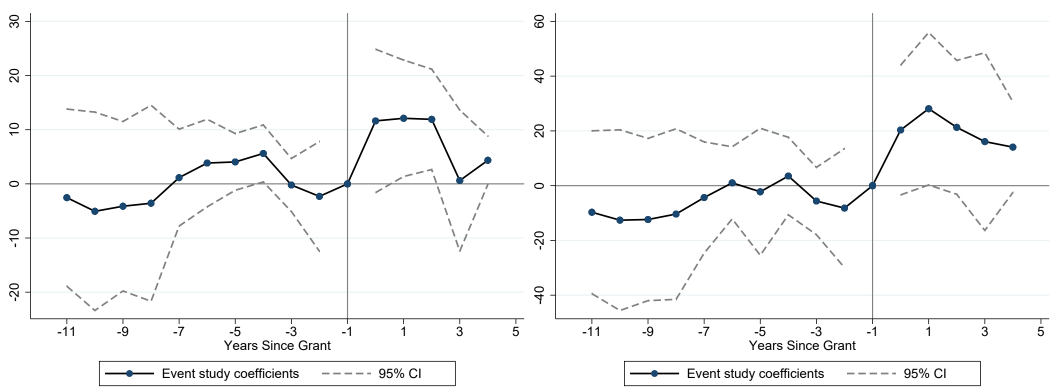
Standard errors clustered by census tract (119 clusters) are in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of SIG-funded interventions on neighborhood composition, as measured by Consumer Finance Protection Bureau mortgage data identifiable at the census tract. Estimates reflect coefficients from a stacked difference-in-difference regression that compares changes in characteristics of tracts zoned for schools treated as part of the 2012 or 2015 SIG cohorts to those of untreated schools. The specification controls for tract-by-stack and year-by-stack fixed effects.

The results in Table 2.10 are consistent with gentrification of SIG neighborhoods following treatment. We explore whether the influx of white, high-income households moving into SIG neighborhoods displaced existing residents using block group counts of court-ordered evictions and eviction filings from the Eviction Lab at Princeton University. Block groups are granular enough that we can mimic the sample restrictions of our capitalization analysis by limiting the sample to block groups that intersect with a half-mile buffer surrounding attendance zone boundaries separating SIG and non-SIG schools (excluding block groups that do not fully lie on the SIG or non-SIG side of the boundary). Figure 2.13 depicts this sample. Because we only observe these data from 2000 to 2016, we do not analyze the 2015 cohort of SIG schools.

Figure 2.5 presents two-way fixed effects event studies for evictions per 1,000 residents and eviction filings per 1,000 residents in panels (a) and (b), respectively, restricting the sample to block groups contiguous to SIG attendance zones. After evolving similarly prior to treatment, block groups just inside SIG attendance zones experience more eviction filings and evictions per 1,000 residents than those just outside SIG attendance zones. Difference-in-differences estimates, presented in Table 2.11, indicate that evictions and eviction filings increased by 6.54 and 23.27 per 1,000 residents in SIG neighborhoods. Relative to the pre-treatment means, these magnitudes represent 34.7% and 61.2% increases, respectively.[23]



(a) Evictions per 1,000

(b) Eviction Filings per 1,000

Figure 2.5: TWFE Event Studies: Evictions

Notes: These figures presents event study estimates of the effect of SIG-funded interventions on evictions using data from the Eviction Lab at Princeton University. We limit the sample to block groups that are at least partially contained within a half-mile buffer around SIG attendance zone boundaries. Estimates reflect coefficients from a two-way fixed effects event study regression that compares changes in characteristics of block groups just inside attendance zones for schools treated as part of the 2012 SIG cohort to those of block groups located just outside of treated attendance zones. The specification controls for block group and year fixed effects. Panels (a) and (b) present estimates of the effect of SIG treatment on the number of court-ordered evictions per 1,000 residents and eviction filings (which may result in a ruling for or against eviction, or a settlement between the landlord and tenant) per 1,000 residents, respectively.

[23]Neither our eviction data nor our home sales panel allow us to observe whether homes where landlords evicted tenants were then sold for owner-occupancy or rented continually. We examine whether the share of owner-occupied units changes in SIG school zones using block group-level data from the 2010 Decennial Census and ACS (2013-2019) in Appendix Figure 2.14 and find no effect of SIG on owner-occupancy rates, although we cannot rule out sizable changes in either direction.

Table 2.11: SIG-Induced Gentrification Increases Evictions

	(1)	(2)
	Evictions per 1,000	Eviction Filings per 1,000
SIG	6.540*	23.27*
	(3.782)	(12.12)
N	608	608

Robust standard errors clustered by block group (38 clusters) in parentheses

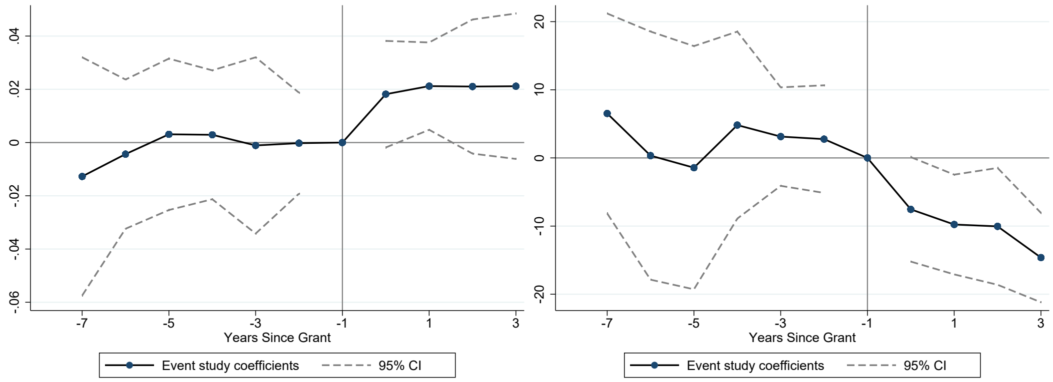
* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of SIG-funded interventions on evictions from a two-way fixed effects DD. We use data on evictions and eviction filings from the Princeton University Eviction Lab (2000 - 2016) and limit the sample to block groups that are at least partially contained within a half-mile buffer around SIG attendance zone boundaries. Because we only observe the data through 2016, we only analyze the 2012 cohort of SIG. Columns (1) and (2) present estimates of the effect of SIG treatment on the number of court-ordered evictions per 1,000 residents and eviction filings (which may result in a ruling for or against eviction, or a settlement between the landlord and tenant) per 1,000 residents, respectively.

To examine how the above sorting affected the overall composition of neighborhoods zoned for SIG schools, we estimate a stacked difference-in-differences with block group-level data from the 2010 Decennial Census and ACS (2013-2019). Because this data only gives us one year of observations prior to treatment for the initial cohort of SIG schools in MNPS, we view this set of results as suggestive evidence and caution against causal interpretation. Appendix Table 2.16 presents coefficients from stacked difference-in-differences estimations. Although noisy, the coefficients are consistent with our mortgage analysis, suggesting that neighborhoods zoned for SIG schools became whiter and wealthier following treatment.

Finally, we examine to what degree changes in neighborhood composition corresponded with changes in classroom composition. Using grade-level enrollment data from NCES, we estimate stacked difference-in-differences and event study specifications that compare changes in enrollment demographics in SIG schools to non-SIG schools in MNPS. Figure 2.6a shows that after evolving in parallel prior to SIG implementation, white share in SIG schools increased by 2 percentage-points following treatment. Consistent with our narrative of displacement,2.6b shows that three years after the start of SIG-funded interventions, 15 fewer nonwhite students were enrolled in a given grade. Table 2.12 presents DD coefficients and confirms the statistical significance of these demographic changes. Relative to the 2010 means, the estimate in Column (1) corresponds to a 19.9% increase in the share of students who are white in SIG schools and that of Column (4) corresponds to a 15% decline in the number of nonwhite students in SIG schools following funding receipt. [24] Columns (2), (3), (5), and (6) show that these changes occurred in both treated elementary and middle schools.

[24] Ideally, we would examine changes in the economic status of student demographics in SIG schools, too. However, measures of student poverty for MNPS schools are not reliable: the average share of students eligible for free or reduced-price lunch in SIG schools in 2010 was a non-sensical 157%.



(a) White Enrollment Share

(b) Nonwhite Enrollment Count

Figure 2.6: Stacked Event Studies: Enrollment

This figure presents coefficients from a stacked event study specification that regresses school-by-grade enrollment demographics on stack-by-school-grade and stack-by-year fixed effects, using school-by-grade enrollment data from NCES. The outcome in panel a) is white enrollment share and in b) is the nonwhite enrollment count.

Table 2.12: School Demographics Reflect Changing Neighborhood Demographics

	(1)	(2)	(3)	(4)	(5)	(6)
	Percent of Students Who Are White			Number of Nonwhite Students		
	All Schools	Elementary	MS	All Schools	Elementary	MS
SIG	0.0221**	0.0249**	0.0269	-12.69**	-7.345**	-18.72
	(0.00948)	(0.0104)	(0.0180)	(6.132)	(3.483)	(11.84)
N	12477	9716	2761	12477	9716	2761

Robust standard errors clustered by school in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This figure presents coefficients from a stacked difference-in-difference specification that regresses school-by-grade enrollment demographics on stack-by-school-grade and stack-by-year fixed effects, using school-by-grade enrollment data from NCES. Column (1) estimates the change in the white enrollment share for all SIG schools relative to non-SIG schools, and Columns (2) and (3) limit the sample to elementary and middle schools respectively. Columns (4), (5), and (6) follow this same pattern for nonwhite enrollment counts.

Overall, estimated effects on mortgage outcomes show that neighborhoods zoned for SIG schools became moderately wealthier and whiter following interventions, estimated effects on evictions indicate that sorting displaced a share of existing residents, and estimated effects on enrollment suggest that school student populations experienced similar demographic changes and displacement. We interpret these results as evidence of SIG funding inducing gentrification of targeted schools and neighborhoods.[25] If instead, the observed gentrification specifically in SIG neighborhoods still would have occurred in the absence of SIG-funded interventions, then these results may be interpreted as evidence that concurrent gentrification can inhibit place-based public goods from reaching their intended recipients.

2.6 Generalizability

To examine the generalizability of our Nashville-based results, we supplement our primary analyses by estimating the effects of SIG-funded interventions on housing prices and neighborhood composition in California. We choose California as our supplementary setting because we are able to observe grant eligibility and receipt for all SIG cohorts in the state and because previous work provides evidence that SIG-funded interventions successfully improved student achievement at low-performing schools in California (Friday, 2021).

2.6.1 Data and Methods

Replicating our Nashville-based analyses in California requires measures of housing prices, neighborhood and school demographics, and displacement linked to school attendance zones. To measure housing prices in neighborhoods across the entire state, we follow previous literature (Bayer et al., 2020) in utilizing tract-level house price indices (HPIs) constructed by the Federal Housing Finance Agency (FHFA). Similarly to Case-Shiller indices (Case and Shiller, 1989), FHFA HPIs are constructed based on changes in the value of houses that have been sold or refinanced multiple times. We again use loan-level mortgage data from the Consumer Financial Protection Bureau, school demographic data (from the California Department of Education), and eviction counts from the Eviction Lab to characterize the sorting response to SIG-funded improvements in school quality.

We observe attendance zone boundaries for neighborhood schools in all districts in California from the School Attendance Boundaries Survey, a mapping of attendance zones for 70,000 schools across 12,000 districts across the country undertaken by NCES and the Census Bureau. School assignment zones in much of California differ from those in Nashville in that they do not perfectly overlap in a 'school pathways' structure (i.e., not all students in a given elementary school will advance to the same middle school). Thus, often a single neighborhood straddles multiple assignment zone boundaries for different school levels. We assign treatment to neighborhoods (Census tracts) defined as the average per-pupil SIG grant amount received by a neighborhood's zoned schools, weighted by the share of the neighborhood's population zoned for each SIG school.

[25] An alternative explanation for our results is that gentrifiers themselves advocated to policymakers to apply for SIG funding for schools in neighborhoods with potential for redevelopment. Selection into treatment based on neighborhood development potential would bias our estimates toward finding results consistent with gentrification that do not actually reflect causal effects of increased school funding. However, the nature of SIG's grant allocation process gave little room for this tape of manipulation. Eligibility for SIG was primarily determined by a school's standardized test scores falling in the bottom five percent of its state. Moreover, nearly all of MNPS schools that were eligible for SIG by this rigid criteria were awarded grants, leaving almost no leeway for selection into grant receipt in our setting. Anecdotally, we find no narrative of parents or commercial groups lobbying to policymakers for specific schools in Davidson County to receive SIG funding in media.

2,687 schools in California were eligible for SIG, 162 of which which received grants averaging $4,077 per-pupil to implement school turnaround interventions across four cohorts (2010-2011, 2011-2012, 2014-2015, and 2016-2017). Table 2.13 presents pre-treatment descriptive statistics and balance tests for SIG-recipient and eligible non-recipient neighborhoods and schools. Much like in Nashville, baseline characteristics of SIG schools and neighborhoods in California differed from those of their eligible non-SIG counterparts. Home values were lower in SIG neighborhoods, and homebuyers in SIG neighborhoods reported lower income and were approximately half as likely to be white and twenty percentage-points more likely to be Hispanic than those in neighborhoods zoned for SIG-eligible schools that did not receive grants. The student populations of treated and SIG-eligible neighborhoods differed less substantially than their home-owning populations, however: student populations at SIG schools had a 9.7 percentage-point smaller share of white students, with this difference evenly accounted for by larger proportions of Black and Hispanic students. Both SIG-recipients and SIG-eligible schools consisted of majority Hispanic populations.

Table 2.13: California Pre-Treatment Descriptive Statistics

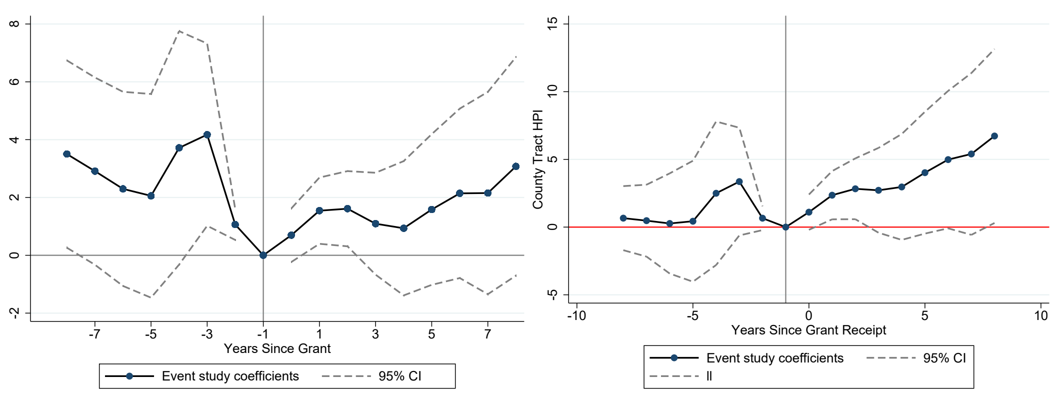
	(1) SIG-Eligible Non-Recipients	(2) SIG-Recipients	(3) Diff
Panel A. Housing Prices			
Tract HPI	348	299	-49***
	(186)	(154)	(3)
Observations	25,587	3,402	28,989
Panel B. Mortgage Characteristics			
Loan Amount	305,906	277,273	-28,632***
	(163,130)	(101,105)	(4,382)
Applicant Income	111,519	91,110	-20,410***
	(76,842)	(38,770)	(2,043)
Share White Buyers	0.450	0.231	-0.219***
	(0.244)	(0.219)	(0.007)
Share Hispanic or Latino Buyers	0.276	0.464	0.189***
	(0.241)	(0.286)	(0.007)
Share Black Buyers	0.031	0.098	0.067***
	(0.054)	(0.149)	(0.002)
Observations	8,694	1,476	10,170
Panel C. Evictions			
Evictions per 1,000	4.237	5.076	0.839*
	(37.441)	(6.289)	(0.468)
Eviction Filings per 1,000	5.042	5.950	0.908
	(46.840)	(7.248)	(0.585)
Observations	40,590	6,800	47,390
Panel D. School Demographics			
Non-Hispanic White Enrollment Share	0.171	0.074	-0.097***
	(0.190)	(0.088)	(0.006)
Non-Hispanic Black Enrollment Share	0.088	0.146	0.058***
	(0.123)	(0.165)	(0.004)
Hispanic Enrollment Share	0.646	0.703	0.057***
	(0.246)	(0.208)	(0.008)
Observations	21,560	1,010	22,570

Notes: This table presents summary statistics and balance tests for pre-SIG neighborhood and school characteristics using data from the Federal Housing Finance Agency (Panel A), the Consumer Financial Protection Bureau (Panel B), the Eviction Lab at Princeton University (Panel C), and the California Department of Education (Panel D). Census tracts in Panels A through C are considered treated if they at least partially fall in an assignment zone for a school that later implemented a SIG-funded turnaround; similarly, control tracts are those that at least partially fall in an assignment zone for a school that was eligible for SIG but never received a grant. P-values for Column 3 represent standard t-tests of mean equality.

To identify the effects of SIG-funded interventions on housing prices and neighborhood composition in California, we estimate a stacked difference-in-differences specification that compares changes in outcomes of neighborhoods zoned for SIG-recipient schools to those zoned for schools that were eligible for, but did not receive, SIG grants. Causal interpretation of these estimates requires the assumption that conditional on grant eligibility, grant receipt was exogenous to trends in neighborhood outcomes. To assess the plausibility of this assumption–specifically, the implication that outcomes in treated and eligible tracts would evolve in parallel prior to treatment–and to trace out potential dynamic treatment effects, we estimate flexible event study specifications.

2.6.2 Results

We first present event study coefficients in Figure 2.7 that represent the changes in housing prices in neighborhoods zoned for SIG-recipient schools associated with a $1,000 increase in per-pupil SIG funding relative to those zoned for eligible non-recipient schools. Panel (a) shows that prior to treatment, housing prices in treated neighborhoods declined relative to those in control neighborhoods. Following treatment, this trend reversed: housing prices in neighborhoods zoned for SIG schools increased relative to prices in neighborhoods zoned for eligible schools. Although the break from trend suggests that SIG-funded interventions had a positive effect on housing prices, the differential pre-trends cast doubt on the identifying assumption of parallel trends between treated and control neighborhoods in the absence of treatment and would bias difference-in-differences estimates downward.



(a) FHFA Housing Price Index

(b) Adjusted for Pre-trends

Figure 2.7: Stacked Event Studies: California Housing Prices

Notes: These figures prevent event study coefficients from stacked event study specifications that compare changes in tract-level housing price indices from the Federal Housing Finance Agency for neighborhoods zoned for schools receiving SIG grants to those zoned for SIG-eligible non-recipients. Treatment in both panels is defined as the average per-pupil SIG grant amount received by a tract's zoned schools, weighted by the share of the tract's population zoned for each SIG school. Coefficients represent the increase in housing prices associated with a $1,000 increase in per-pupil SIG funding. Panel (a) is a standard event study, and standard errors are clustered by county. Panel (b) adjusts for pre-trends following a process outlined by Goodman-Bacon (2021c), first estimating a pre-trend in housing prices for treated and control groups and removing this trend from the full panel prior to estimating the event study. To account for variation from the first-step detrending estimation, we construct two-step bootstrap standard errors clustered at the county level.

We account for bias from group-specific linear pre-trends by using a procedure proposed by Goodman-Bacon (2021c), estimating a pre-trend in housing prices for treated and control groups and removing this trend from the full panel.[26] The resulting event study and difference-in-differences estimations on the pre-trend-adjusted outcome variable are robust to linear trends and time-varying treatment effects. To interpret the adjusted estimations as causal effects of SIG-funded interventions, we must assume that housing prices would continue to decline linearly in treated neighborhoods relative to control neighborhoods in the absence of treatment. Because this is stronger than the identifying assumption of the standard difference-in-differences estimator, we cautiously interpret these results as suggestive.

Panel (b) of Figure 2.7b presents the housing price index event study adjusted for group-specific linear pre-trends. Following grant receipt, home prices steadily increase in neighborhoods zoned for schools implementing SIG-funded interventions. Treatment effect growth over time mirrors the response found in Nashville. The stacked difference-in-differences specification, presented in Table 2.14, summarizes this dynamic response as a 2.6 unit increase in HPI per $1,000 in SIG funding. This coefficient corresponds to a 3.6% housing price increase in the average SIG-zoned neighborhood upon receipt of the average grant of $4,077. Specifications using mortgage data from the CFPB can be used to construct a specific dollar amount increase in price, allowing for an Oates (1969) efficiency test. As shown in Table 2.14, the average mortgage increased by 3.0 log points (3.0%) for every $1,000 in SIG funding. Based on the average pre-treatment loan size in SIG-zoned neighborhoods of $277,273, this represents over an $8,000 increase in housing prices. The 8:1 return on SIG investments is even larger than the corresponding return estimated in Nashville.

Table 2.14: Effect of SIG on Housing Prices in California

	(1)	(2)	(3)	(4)	5)	(6)	(7)	(8)
	HPI	ln(Loan Amount)	ln(App. Income)	Share White Buyers	Evictions per 1,000	Filings per 1,000	White Enr. Share	Hispanic Enr. Share
SIG	2.634**	0.030**	0.035***	0.007***	0.280	0.277	0.0005	-0.0027***
	(1.098)	(0.014)	(0.009)	(0.003)	(0.171)	(0.203)	(0.0009)	(0.0010)
N	112,119	65,952	65,952	65,952	123,744	123,744	83,847	83,847

Two-step bootstrap standard errors clustered at the county level for Columns (1) through (6) and at the school district level for Columns (7) and (8) are in parenthesis

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates from stacked pre-trend-adjusted difference-in-differences specifications that compare changes in outcomes of schools and neighborhoods receiving SIG grants to SIG-eligible non-recipients. Treatment across Columns (1) through (6) (tract-level outcomes) is defined as the average per-pupil SIG grant amount received by a tract's zoned schools, weighted by the share of the tract's population zoned for each SIG school; treatment in Column (7) (a school-level outcome) is defined as the per-pupil SIG grant received by a school. Coefficients across all specifications represent the change in outcomes associated with a $1,000 increase in per-pupil funding. All specifications adjust for group-specific linear pre-trends following a process outlined by Goodman-Bacon (2021c), first estimating a pre-trend in outcomes for treated and control groups and removing this trend from the full panel prior to estimating the difference-in-differences. To account for variation from the first-step detrending estimation, we construct two-step bootstrap standard errors. Column (1) uses tract-level housing price indexes from the Federal Housing Finance Agency. Columns (2) through (4) use tract-level mortgage data from the consumer Financial Protection Bureau. Columns (5) and (6) use tract-level eviction counts from the Eviction Lab at Princeton University. Finally, Columns (7) and (8) use school enrollment data from the California Department of Education.

Just as in Nashville, the capitalization of SIG-funded interventions into housing prices in California may

[26] To account for additional variance generated by the first-step estimation, we construct standard errors utilizing a two-step bootstrap procedure clustered at the county or school-district level.

represent substantial neighborhood change and distort who can access newly improved schools. We characterize the sorting response to SIG in California by estimating pre-trend-adjusted stacked difference-in-differences specifications for the effect of SIG-funded interventions on homebuyer characteristics, displacement, and school demographics and present coefficients in Table 2.14.[27] Our results suggest income-based sorting of even greater magnitude than in Nashville: reported income of homebuyers in SIG neighborhoods increased by 15.3% in response to the average grant. Moreover, although less pronounced than in Nashville, we still find evidence of racial sorting and displacement from neighborhoods zoned for newly improved schools. Overall, Table 2.14 supports the notion that the effects of SIG-funded interventions on housing prices and neighborhood composition were not unique to Nashville.

2.7 Discussion & Conclusion

This paper examines whether household sorting in response to School Improvement Grant funding inhibits spending from reaching targeted students. Estimating a boundary-discontinuity difference-in-differences model on property sales data from 2000 to 2019, we find that home values in SIG school zones increased by 10.5% relative to houses just across attendance zone boundaries in response to SIG-funded interventions. This increase in price coincided with wealthier, whiter households moving into SIG attendance zones and schools and the displacement of disadvantaged residents and students, as evident by a 35% increase in evictions and a 15% decline in the number of nonwhite students in SIG schools. Supplementary analyses in California suggest that this phenomenon was not unique to Nashville.

Although previous research has found generally positive effects of school turnaround programs on student achievement (Friday, 2021; Sun et al., 2020; Carlson and Lavertu, 2018; Sun et al., 2017), we provide direct evidence that households observe and value improvements in school quality from SIG-funded interventions. Estimates from our preferred specification imply that households were willing to pay $13,148 to move into a SIG school zone, more than a 3:1 return on the average per-pupil total grant award of $3,817.57 to MNPS schools.[28] [29] In the framework of Oates (1969), a dollar of outside funding (i.e., that which does not need to be paid by local taxes) toward a public good should increase housing prices by more than $1 if the public good is under provided. Our estimated willingness to pay of over $3 for a dollar in SIG funding is consisent with the hypothesis program was an efficient use of resources in Nashville, and the replication of this result in California is consistent with suboptimal funding of education in the nation's lowest-performing schools more generally. We contribute to literature on the capitalization of school quality by examining willingness-to-pay for funding targeting schools specifically at the bottom of the achievement distribution, where additional spending may be especially impactful.

We characterize who gained access to newly improved schools by examining characteristics of households

[27] The corresponding event studies are presented in Appendix Figures 2.15 through 2.17.

[28] $13,148 = 10.5% (i.e., 10.0 log points) × $125,216, the mean pre-treatment house price in SIG school zones

[29] Although our estimated price effects fall on the larger end of estimates from the capitalization literature, they are qualitatively similar to the magnitudes of estimates of longstanding capitalization of school quality (Kane et al., 2006) and sorting Billings et al. (2017) found in Charlotte, NC, a comparable setting to Nashville.

moving into neighborhoods zoned for schools implementing SIG-funded interventions. Following grant receipt, higher-income, white households differentially moved into previously majority-minority SIG school zones, and the share of students who are white at SIG schools increased by nearly 20%. Billings et al. (2017) similarly find gentrification of neighborhoods that suddenly receive access to higher quality schools in Charlotte, North Carolina,[30] and Bayer et al. (2020) find that school poverty rates decline by 0.21% in response to a 1% increase in school spending. These results and this paper's findings support the notion that households may sort across neighborhoods to access newly improved schools, a phenomenon that requires consideration when designing targeted K-12 funding and assignment policies.

The extent to which displaced households and students are "losers" of SIG-funded interventions depends on the quality of neighborhoods and schools to which they relocate. Because SIG targeted the country's lowest-performing schools in Nashville, with only *student* displacement, displaced students would mechanically move into schools with test scores at least as high as their original schools prior to SIG funding receipt. However, because SIG funding was largely used as 'hazard pay' to attract higher-quality teachers to targeted schools, the realized quality of local non-SIG schools populated by displaced students could be negatively affected by *teacher* poaching (Kho et al., 2020). Future work with a general equilibrium analysis of SIG's effect on teacher labor markets and the quality of schooling experienced by displaced students would fill this gap in policy evaluations of SIG and other K-12 funding interventions.

More broadly, our sorting results relate to the literature on public good provision and racial and income-based segregation. Moderate integration following SIG funding receipt is consistent the results of Diamond and McQuade (2019), whom find that place-based investments via the form of Low Income Housing Tax Credit-funded developments in low-income, high-minority neighborhoods raise house prices and increase neighborhood socioeconomic diversity. We cannot speak to whether the observed moderate integration into neighborhoods experiencing improved school quality will persist, and previous literature lends support to either lasting integration or a reversal. Card et al. (2011) find no evidence of minority flight when minority-majority neighborhoods experience small influxes of white neighbors; however, Banzhaf and Walsh (2013) find that race-based sorting dominates income-based sorting based on public goods when differences in public good quality across neighborhoods diminish.

Finally, our results contribute to the literature on the nature of gentrification of low-income neighborhoods, a central question of which is whether gentrification helps or harms (predominantly minority) existing neighborhood residents. McKinnish et al. (2010) find no evidence of displacement of low-education or minority residents of gentrifying neighborhoods during the 1990s, despite large-scale in-migration of college-educated whites that increased housing prices and neighborhood income. We provide evidence that gentrification following SIG funding receipt displaced a non-negligible share of existing residents through increased

[30]The authors find that average home price and household income in the "highest-quality" neighborhoods that qualify for priority in enrollment lotteries for oversubscribed schools increase by 8.4 and 13.1 log points, respectively. These results are qualitatively similar to the housing price and income effects we find (10.0 and 8.6 log point respective increases), and come from a comparable setting to ours, in regards to both the similarity of Nashville and Charlotte's housing markets and that both our paper and Billings et al. (2017) examine policies that affect attendance zones for low-performing schools.

evictions. This phenomenon demonstrates a limitation of place-based public good provision: sorting may displace the initially targeted population. Normative evaluations of policies that directly or indirectly affect housing markets will benefit from further research on this dynamic.

2.8 Appendix

Table 2.15: SIG Improves the Lower-Tail of Student Achievement

	(1)	(2)	(3)	(4)
	RLA	Grade x Year FEs	Math	Grade x Year FEs
SIG	-4.034	-3.933	-4.239	-2.272
	(3.017)	(2.995)	(8.903)	(9.088)
N	940	940	940	940

Robust standard errors clustered by school are in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents difference-in-difference estimates of the test score effects of SIG for the 2012 cohort in MNPS, using school-by-grade achievement data from the Tennessee Department of Education from the 2010-2011 through 2014-2015 school years. The dependent variable is the share of students scoring “Below Basic”, the lowest of TDOE s four achievement benchmarks, in reading (columns 1 and 2) and math (columbs 3 and 4). Columns 1 and 3 include school-by-grade and year fixed effects, while columns 2 and 4 include school-by-grade and grade-by-year fixed effects. We cannot extend the panel beyond 2014-2015, and thus cannot examine test score effects of the 2015 cohort of SIG, because of a change in testing regimes that makes comparisons of test scores before and after the transition uninformative.

Table 2.16: Block Group DDs: ACS Resident Characteristics

	(1)	(2)	(3)	(4)
	ln Population	ln White Pop	Median Household Income	Median Gross Rent
SIG	-0.0300	0.130	509.1	7.891
	(0.0357)	(0.119)	(1553.7)	(15.16)
N	506	506	506	506

Standard errors clustered by census block group (52 clusters) are in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the effect of SIG-funded interventions on neighborhood composition using census block group-level data from the 2010 Census and American Communities Survey (2013-2019). We limit the sample to block groups that are at least partially contained within a half-mile buffer around SIG attendance zone boundaries. Estimates reflect coefficients from a stacked difference-in-difference regression that compares changes in characteristics of block groups just inside attendance zones for schools treated as part of the 2012 or 2015 SIG cohorts to those of block groups located just outside of treated attendance zones. The specification controls for block group-by-stack and year-by-stack fixed effects.

Table 2.17: Falsification Exercise: Shifting SIG Boundaries

	(1) 1 Mile	(2) 3 Miles	(3) 5 Miles
A. West			
Placebo SIG	0.130	0.0304	-0.0611
	(0.0838)	(0.121)	(0.120)
N	23198	16241	13723
B. East			
Placebo SIG	0.110	0.0129	-0.0441
	(0.0726)	(0.0844)	(0.126)
N	22781	17549	17727
C. North			
Placebo SIG	-0.0584	0.0891	-0.0172
	(0.0892)	(0.0998)	(0.0452)
N	20857	15444	12428
D. South			
Placebo SIG	0.0101	0.231**	0.0406
	(0.0686)	(0.0959)	(0.0600)
N	26767	21217	17361

Robust standard errors clustered by boundary segment-by-post are in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents stacked difference-in-differences estimates from a set of falsification exercises that compare changes in housing prices within a half mile on either side of placebo SIG attendance zone boundaries. The placebo boundaries are constructed by moving the real attendance zone boundaries 1, 3, and 5 miles in each cardinal direction. All estimations include school zone-by-stack and stack-by-year fixed effects and control for observable home characteristics, corresponding to our preferred specification from Table 2.7. The sample is limited to single family homes.

Table 2.18: Capitalization Stacked DD: Results are Robust to “Bad Controls”

	(1)	(2)	(3)
	.5 mi	+ Controls	+ Bad Controls
SIG	0.114^{**}	0.107^{**}	0.0979^{**}
	(0.0503)	(0.0429)	(0.0372)
N	10511	10511	10511

Robust standard errors clustered by boundary segment-by-post (46 clusters) are in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of β from equation 2.3, a stacked difference-in-difference specification that compares changes in the logged sale price of homes sold just on either side of SIG attendance zone boundaries. All estimations include school zone-by-stack and stack-by-year fixed effects. The analysis sample is restricted to single-family home sales between 2007 through 2017. Column (1) limits the sample to homes within a half-mile of a SIG attendance zone boundary and includes boundary segment fixed effects. Column (2) adds controls for observable home characteristics. Columns (3) adds two time-varying neighborhood-level controls from Consumer Financial Protection Bureau mortgage data: the natural log of the average reported income of homebuyers and the share of homebuyers that are white.

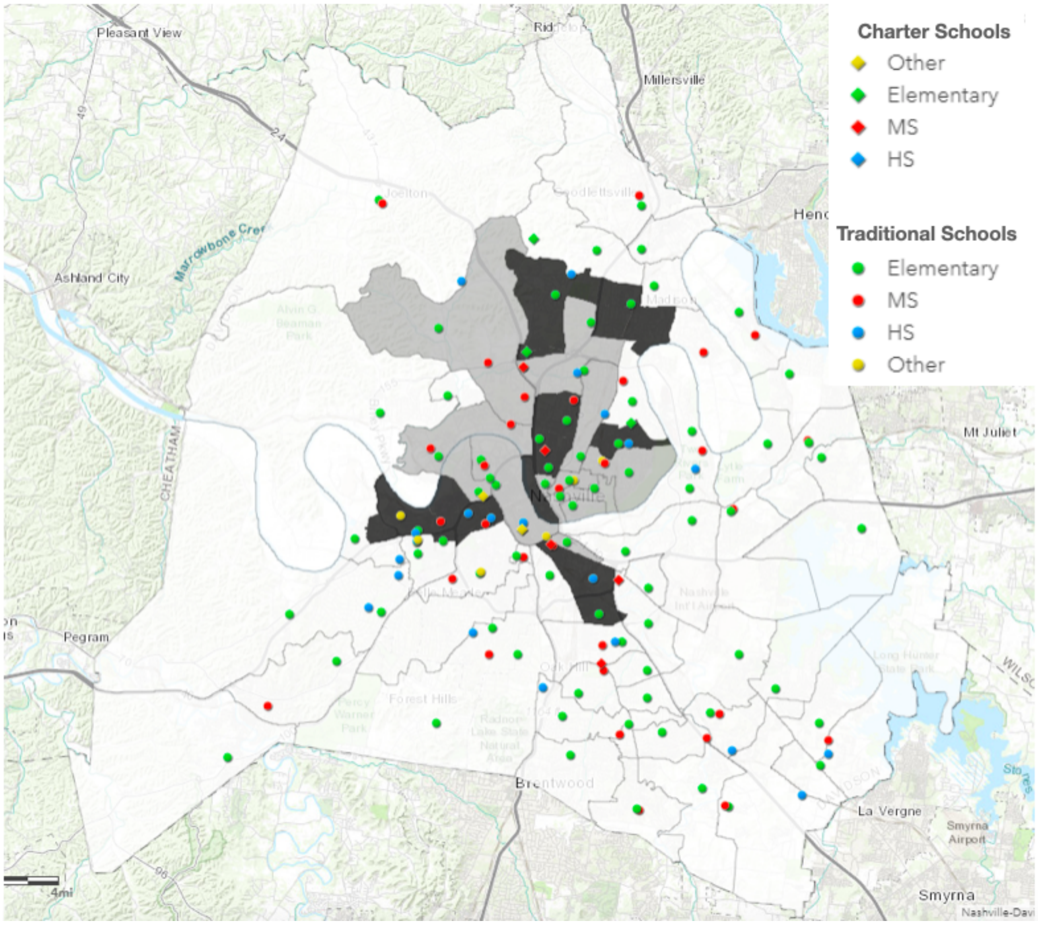


Figure 2.8: MNPS School Zones and School Locations

Notes: This figure maps MNPS school zones and school locations. Attendance zones for schools receiving school improvement grants in the 2012-2013 and 2015-2016 school years are shaded in gray and Black, respectively. Elementary, middle, and high schools are marked in green, red, and blue (schools with multiple levels are marked in yellow). Adhering to these coloring's, charter schools are denoted by diamonds, while traditional schools are denoted by circles. Each attendance zone corresponds to a pathway of assigned elementary, middle, and high schools.

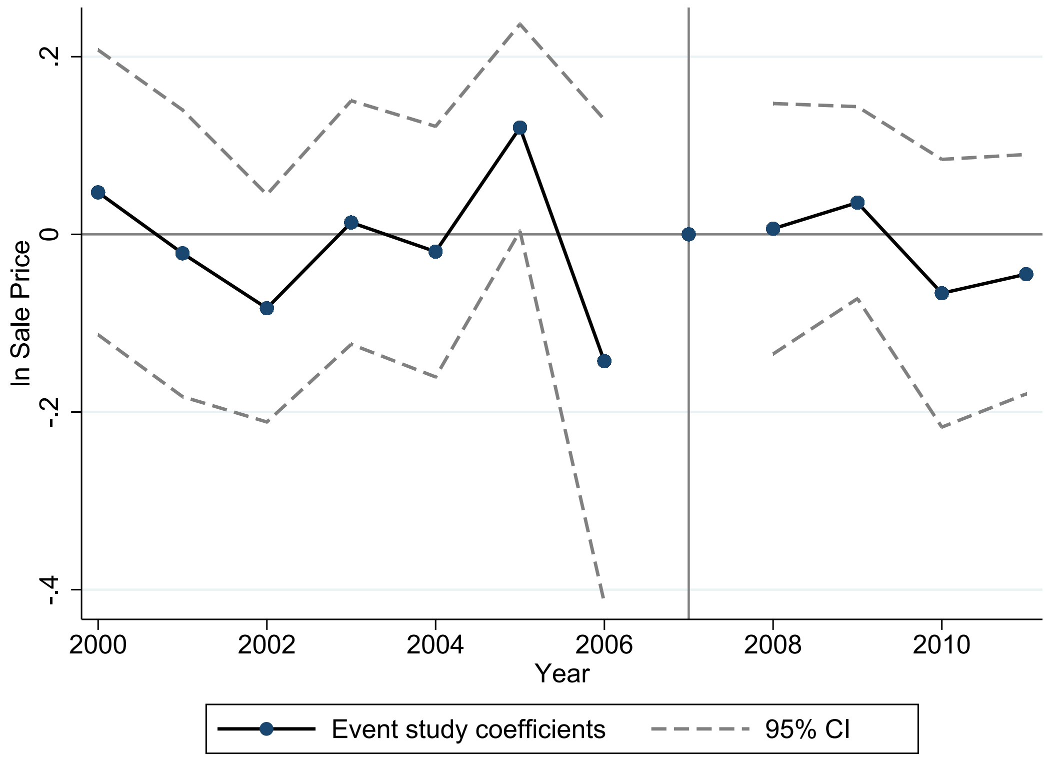


Figure 2.9: Falsification Test: Housing Prices Along SIG Boundaries During The Crash

Notes: This figure plots coefficients a two-way fixed effects event study specification that compares changes in the natural log of sale price (in 2010 $) through the 2007 Housing Crash of homes located on either side of attendance zone boundaries of schools that later implemented SIG-funded interventions. The sample is limited to single family homes and duplexes located within a half-mile of attendance zone boundaries separating treated and untreated school zones. Dashed lines represent the 95% confidence interval of estimates based off of robust standard errors clustered by boundary segment-by-post.

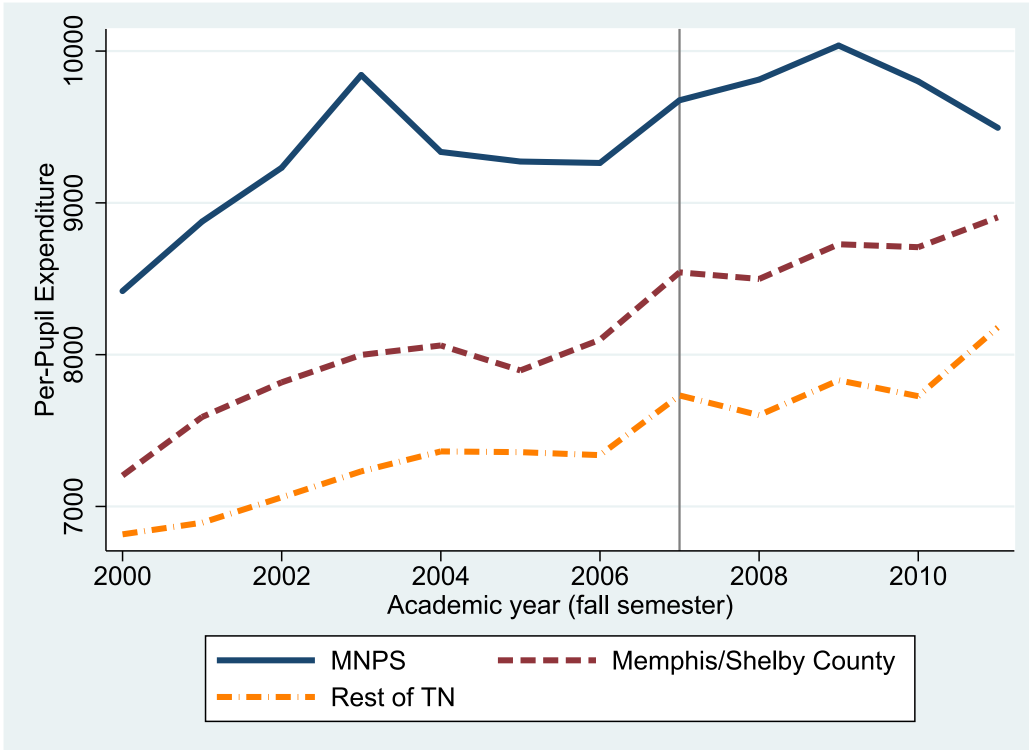
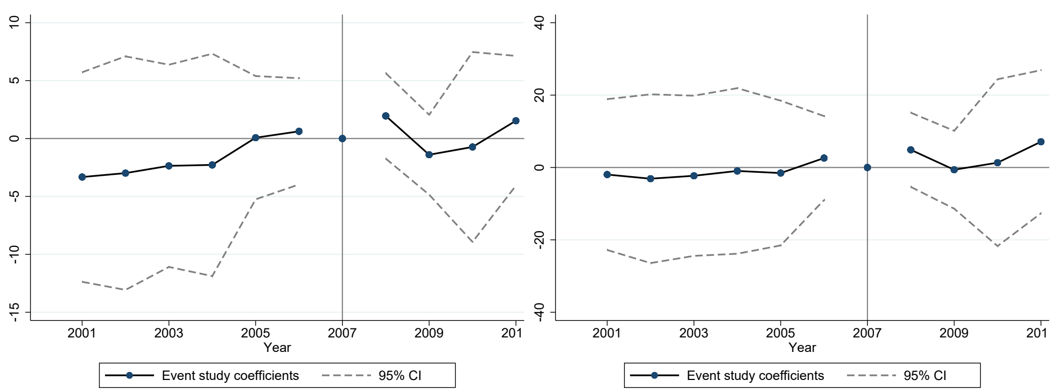


Figure 2.10: School Spending through the Great Recession in MNPS and TN

Notes: This figure plots annual per-pupil expenditure (2010$) in MNPS, Shelby County Schools, and the rest of Tennessee from 2000 to 2011. The vertical line marks the start of the housing crash and ensuing recession.



(a) Evictions per 1,000

(b) Eviction Filings per 1,000

Figure 2.11: Falsification Test: Evictions Along SIG Boundaries During The Crash

Notes: This figure plots coefficients a two-way fixed effects event study specification that compares changes in eviction rates through the 2007 Housing Crash and Great Recession in neighborhoods on either side of attendance zone boundaries of schools that later implemented SIG-funded interventions. We limit the sample to block groups that are at least partially contained within a half-mile buffer around SIG attendance zone boundaries. The specification controls for block group and year fixed effects. Panels (a) and (b) present estimates of the effect of the crash and reccession on the number of court-ordered evictions per 1,000 residents and eviction filings (which may result in a ruling for or against eviction, or a settlement between the landlord and tenant) per 1,000 residents, respectively.

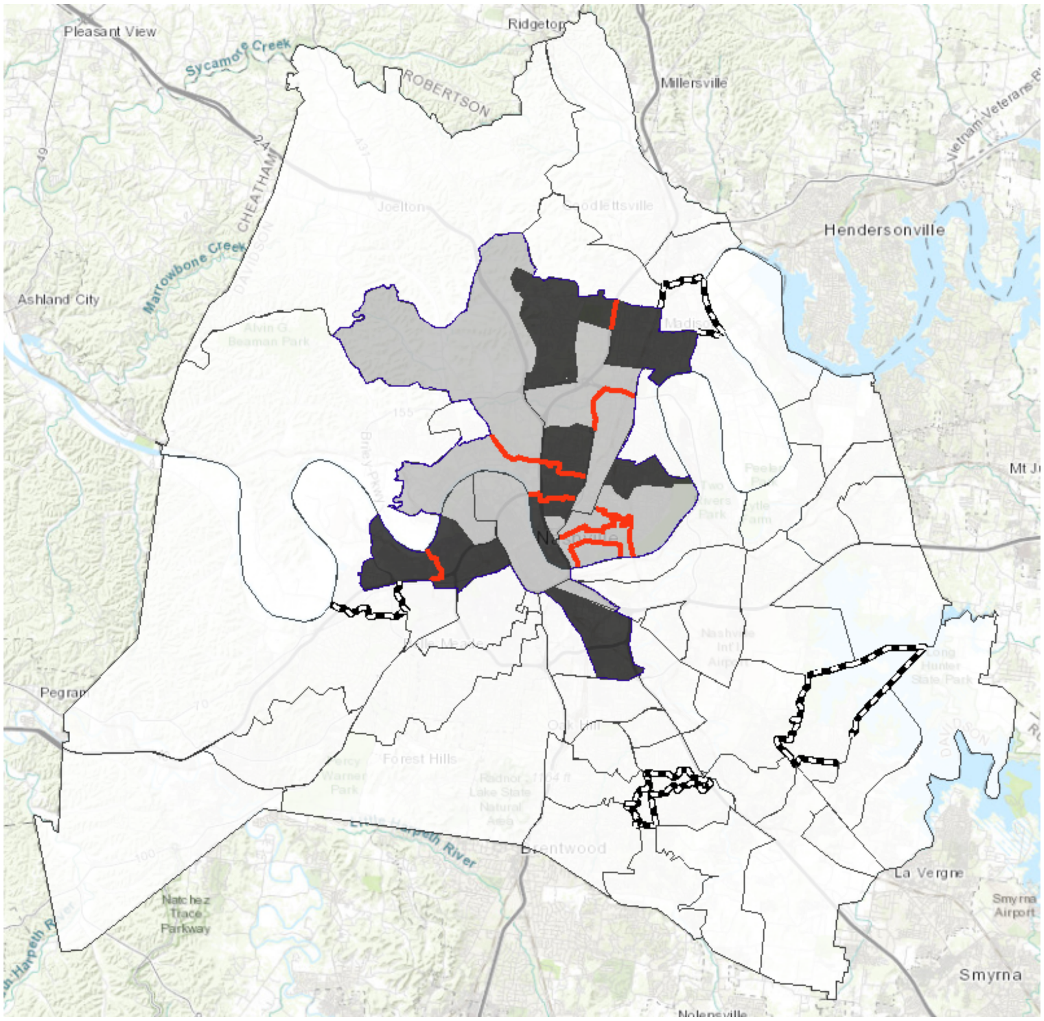


Figure 2.12: Real and Placebo SIG Treatment

Notes: This figure imposes attendance zone boundaries used for falsification exercises over a SIG treatment map of MNPS attendance zones. Each attendance zone corresponds to a pathway of assigned elementary, middle, and high schools. Solid boundary segments separate untreated elementary schools that fall within the attendance zone for the same treated middle or high school. Dashed boundary segments denote attendance zone boundaries for schools that were eligible for SIG but did not receive grants to implement turnaround programs. Results from the falsification exercises are presented in Table 2.9.

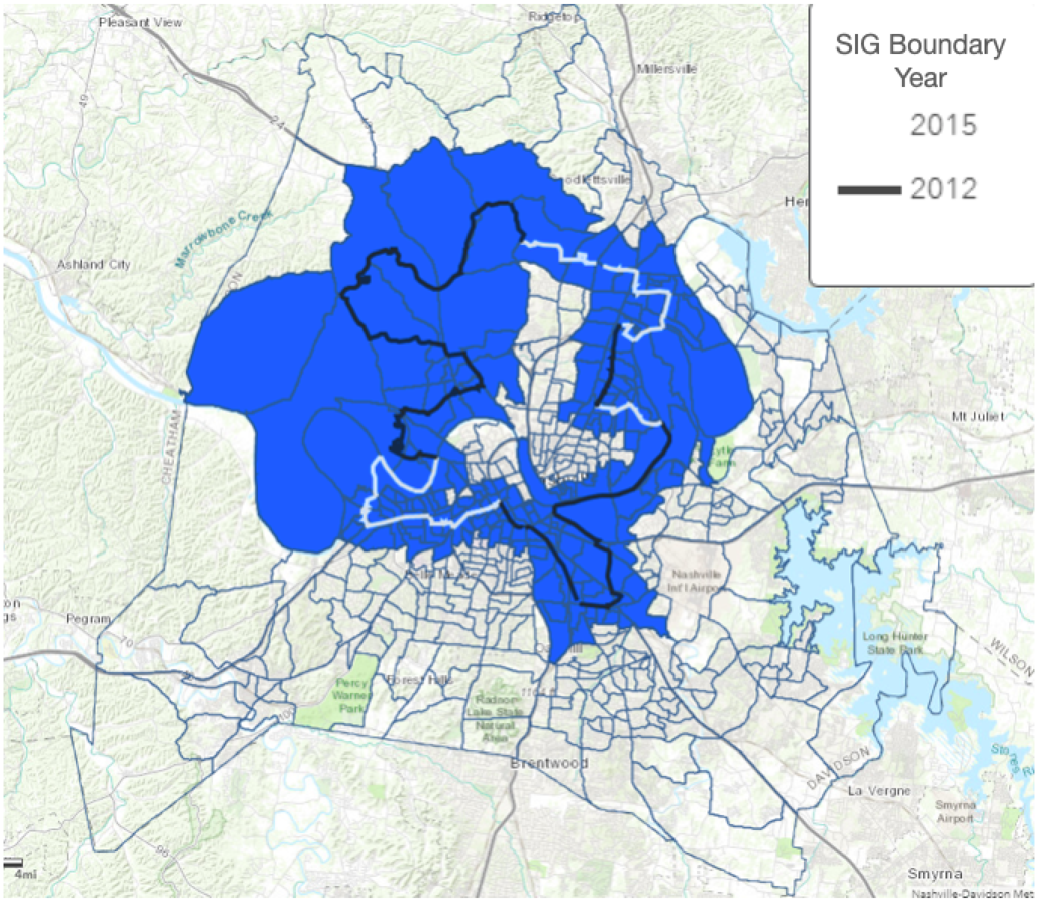


Figure 2.13: Contiguous Census Block Groups

Notes: This figure portrays the sample of "contiguous" Census block groups for estimations of the affect of SIG-funded interventions on neighborhood composition that use block group-level data. The sample consists of block groups that intersection with a half-mile buffer surrounding the boundary separating attendance zones for SIG and non-SIG schools. Block groups that do not completely fall on one side or the other of the boundary (i.e., those that are partially treated) are tossed out.

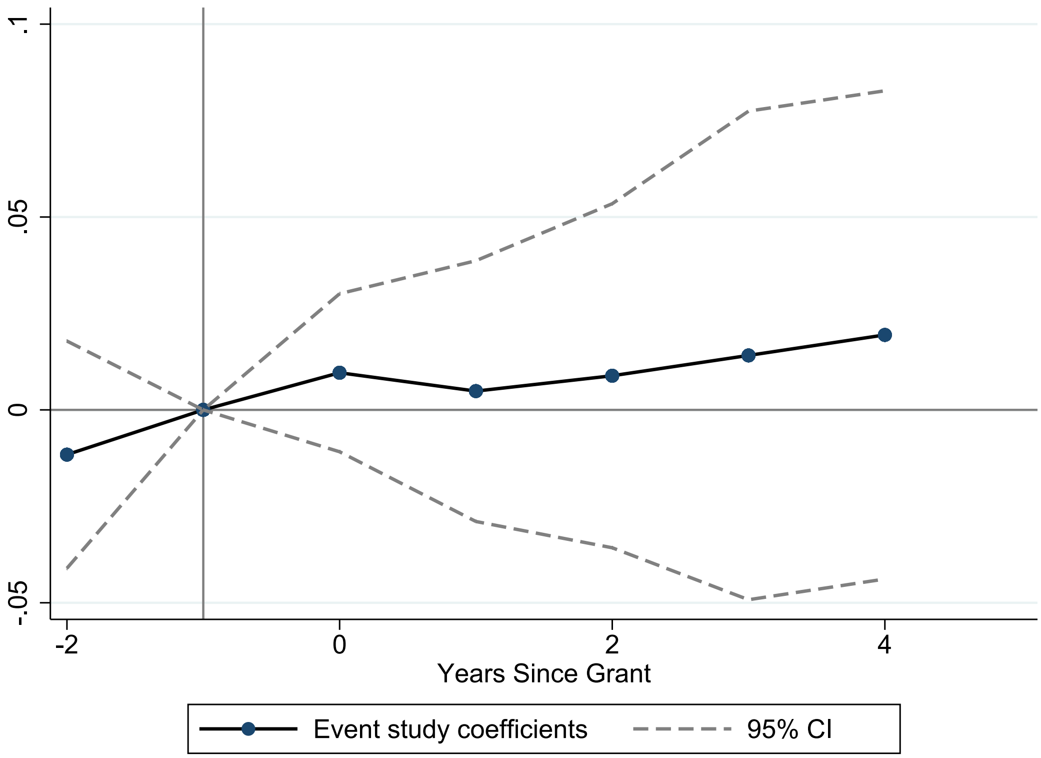
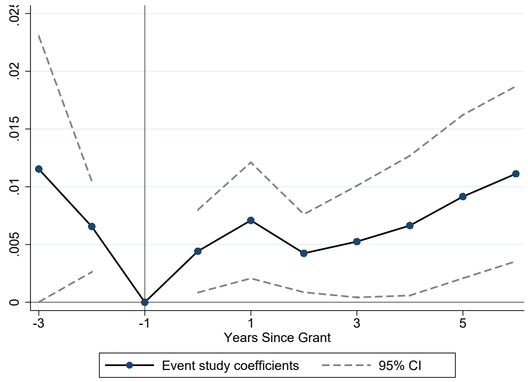
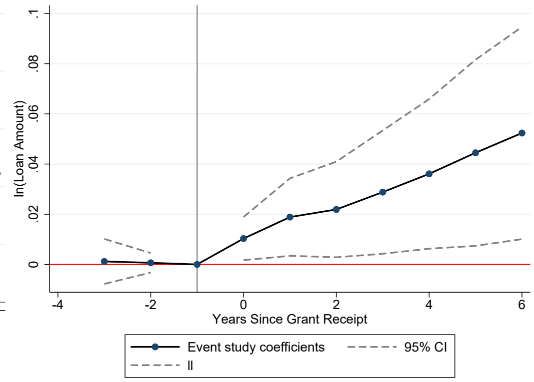


Figure 2.14: Stacked Event Study: Owner-Occupancy Rate

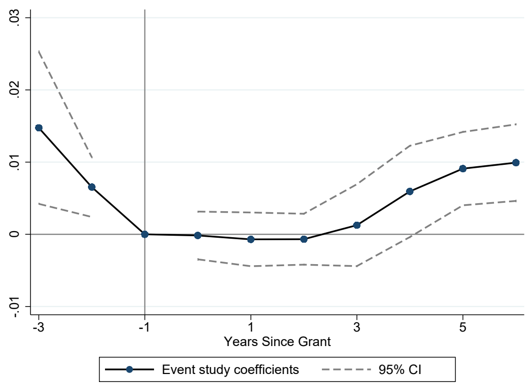
Notes: This figure presents estimates from a stacked event study specification of the effect of SIG-funded interventions on owner-occupancy rates using Census block group-level data from the 2010 Census and American Communities Survey (2013-2019). We limit the sample to block groups that are at least partially contained within a half-mile buffer around SIG attendance zone boundaries. Estimates reflect coefficients from a stacked event study regression that compares changes in characteristics of block groups just inside attendance zones for schools treated as part of the 2012 or 2015 SIG cohorts to those of block groups located just outside of treated attendance zones. The specification controls for block group-by-stack and year-by-stack fixed effects. Dashed lines represent the 95% confidence interval of estimates based off of robust standard errors clustered by boundary segment-by-post.



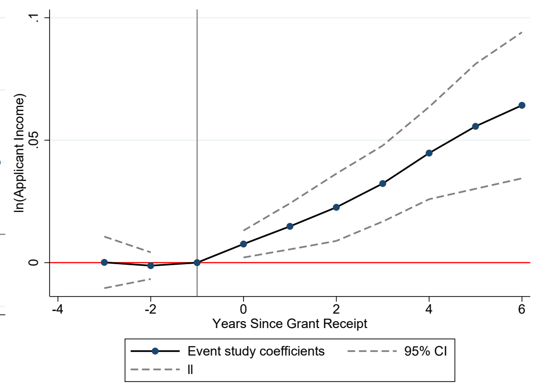
(a) ln(Loan Amount)



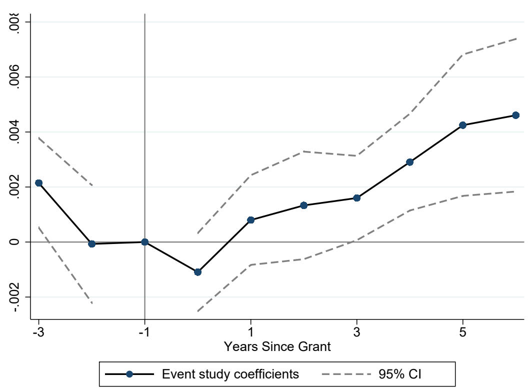
(b) Adjusted for Pre-trends



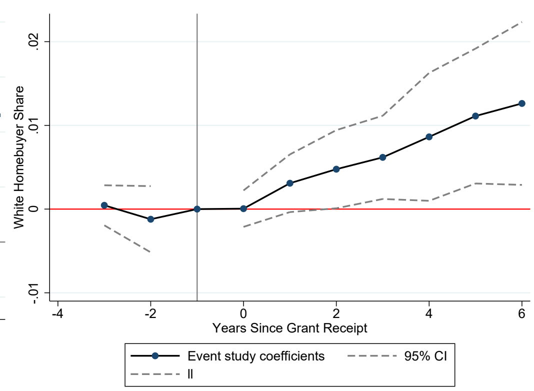
(c) ln(Applicant Income)



(d) Adjusted for Pre-trends



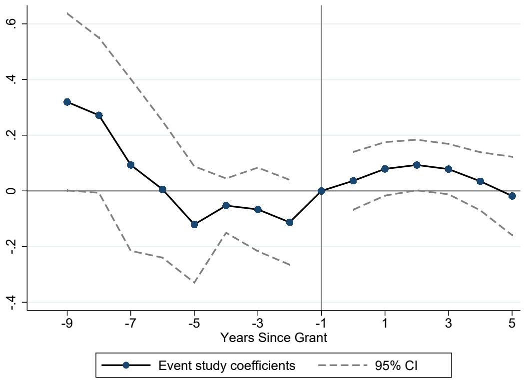
(e) Share White Buyers



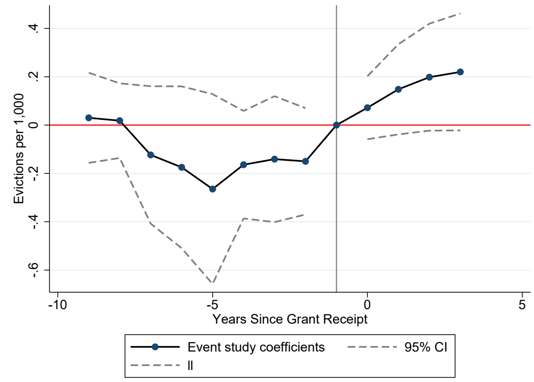
(f) Adjusted for Pre-trends

Figure 2.15: Stacked Event Studies: California Homebuyer Characteristics

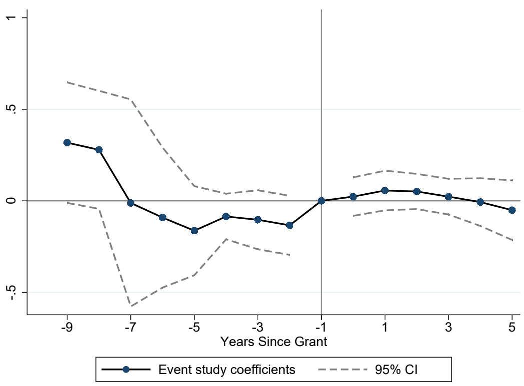
Notes: These figures prevent event study coefficients from stacked event study specifications that compare changes in outcomes mortgages for homes in neighborhoods receiving SIG grants to those in SIG-eligible non-recipient neighborhoods, using data from the Consumer Financial Protection Bureau. Treatment in both panels is defined as the average per-pupil SIG grant amount received by a tract's zoned schools, weighted by the share of the tract's population zoned for each SIG school. Coefficients represent the changes in outcomes associated with a $1,000 increase in per-pupil SIG funding. Panels (a), (c), and (e) are standard event studies, and standard errors are clustered by county. Panels (b), (d), and (f) adjust for pre-trends following a process outlined by Goodman-Bacon (2021c), first estimating a pre-trend in mortgage characteristcs for treated and control groups and removing this trend from the full panel prior to estimating the event study. To account for variation from the first-step detrending estimation, we construct two-step bootstrap standard errors at the county level.



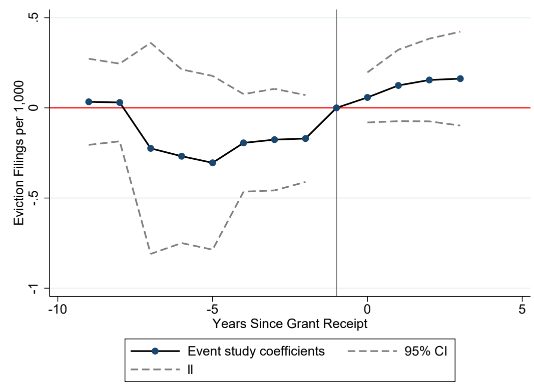
(a) Evictions per 1,000



(b) Adjusted for Pre-trends



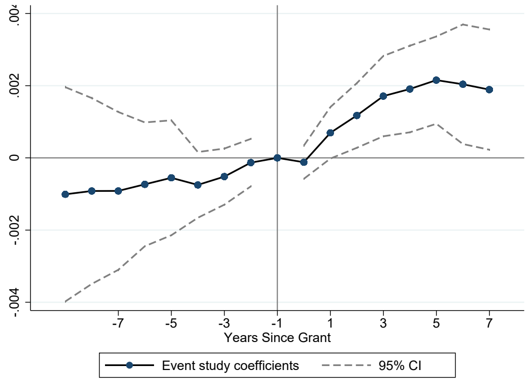
(c) Eviction Filings per 1,000



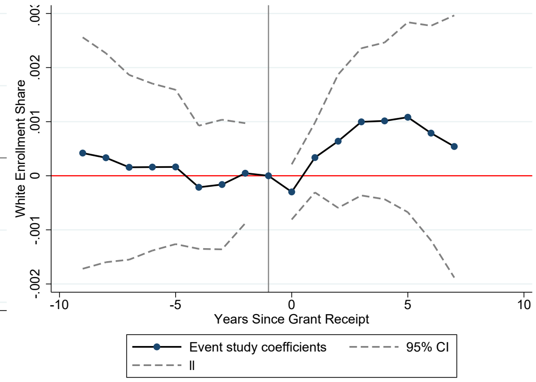
(d) Adjusted for Pre-trends

Figure 2.16: Stacked Event Studies: California Evictions

Notes: These figures prevent event study coefficients from stacked event study specifications that compare changes in outcomes of neighborhoods receiving SIG grants to those in SIG-eligible non-recipient neighborhoods, using data from the Eviction Lab at Princeton University. Treatment in both panels is defined as the average per-pupil SIG grant amount received by a tract's zoned schools, weighted by the share of the tract's population zoned for each SIG school. Coefficients represent the changes in outcomes associated with a $1,000 increase in per-pupil SIG funding. Panels (a), and (c) are standard event studies, and standard errors are clustered by county. Panels (b), and (d) adjust for pre-trends following a process outlined by Goodman-Bacon (2021c), first estimating a pre-trend in eviction outcomes for treated and control groups and removing this trend from the full panel prior to estimating the event study. To account for variation from the first-step detrending estimation, we construct two-step bootstrap standard errors at the county level.



(a) White Enrollment Share



(b) Adjusted for Pre-trends

Figure 2.17: Stacked Event Studies: California School Demographics

Notes: These figures prevent event study coefficients from stacked event study specifications that compare changes in student demographics of schools receiving SIG grants to SIG-eligible non-recipients. Treatment in both panels is defined as the per-pupil SIG grant amount received by a school. Coefficients represent the increase in the white share of students associated with a $1,000 increase in per-pupil SIG funding. Panel (a) is a standard event study, and standard errors are clustered by school district. Panel (b) adjusts for pre-trends following a process outlined by Goodman-Bacon (2021c), first estimating a pre-trend in enrollment demographics for treated and control groups and removing this trend from the full panel prior to estimating the event study. To account for variation from the first-step detrending estimation, we construct two-step bootstrap standard errors at the school district level.

CHAPTER 3

Rivers, Lakes and Revenue Streams: The Heterogeneous Effects of Clean Water Act Grants on Local Spending

This chapter is adapted from "Rivers, Lakes and Revenue Streams: The Heterogeneous Effects of Clean Water Act Grants on Local Spending" published in *Journal of Public Economics* and was reproduced with the permission of the publisher and my co-author Patrick Flynn (Flynn and Smith, 2022).

3.1 Introduction

From 1972 to 1988, the Clean Water Act funded $153 billion (in 2014 dollars) in grants to municipal governments for capital upgrades to wastewater treatment facilities. While CWA grants caused significant reductions in pollution and increases in housing prices (Keiser and Shapiro, 2019), as well as improvements in infant health (Flynn and Marcus, 2021), analyses of the CWA consistently estimate benefit/cost ratios below one (Keiser et al., 2019) with each grant dollar generating an average return of $0.45 (Keiser and Shapiro, 2019; Flynn and Marcus, 2021). Intergovernmental transfers like CWA grants account for around one-fifth of all municipal government revenues (Urban-Brookings Tax Policy Center, 2017), but donor governments often have little control over how receiving governments adjust spending in response to grant receipt. In most cases, having a high rate of "grant pass-through", the increase in targeted expenditure induced by a dollar of grant revenue, is a necessary condition for grant programs to accomplish their policy goals.

This paper uses recently developed difference-in-differences methods (Deshpande and Li, 2019; Callaway and Sant'Anna, 2020b; **?**) to show that, given previously quantified benefits of the CWA, accounting for low pass-through can lead to favorable benefit to cost ratios of CWA-funded wastewater capital upgrades. Our analysis combines data on every CWA grant with spending data from 112 municipal governments. We compare changes in spending in newly treated municipalities to changes in spending in municipalities that will be treated in future periods to show that, on average, each dollar of grant revenue is associated with a $0.45 increase in sewerage capital spending. Dividing previously estimated benefit to cost ratios of CWA grants from Keiser and Shapiro (2019) and Flynn and Marcus (2021) by this pass-through estimate suggests that each CWA grant dollar that municipalities spent on sewerage capital generated an average return of $1.01.

Along with funding grants, the Act imposed new capital standards on all wastewater treatment facilities in the United States. Municipalities could use grant funds to offset the new costs that the CWA's capital mandate imposed on them, but many CWA grants were substantially larger than the cost of coming into compliance with the mandate. This motivates an examination of how pass-through relates to mandated costs. Using data from an EPA engineering assessment of wastewater capital upgrade costs, we divide our sample of municipalities into those that received grants that were approximately equal to mandated costs, and those that received grants that were larger than mandated costs. We show that grants caused a dollar-for-dollar increase in sewerage capital spending in municipalities that received grants that were approximately equal to mandated costs, but in municipalities that received grants that were larger than mandated costs, sewerage capital spending only increased by 27 cents for every dollar of grant funding. The EPA also distributed many CWA grants to municipalities with wastewater treatment facilities that were already in compliance with the CWA's capital mandate. The mandate imposed no costs on these "compliant" municipalities, and they only increased sewerage capital spending by 3 cents for every dollar of grant funding.

The heterogeneous effects of CWA grants on local spending inform the design of intergovernmental grant programs and our understanding of the Clean Water Act. By providing evidence of full pass-through in municipalities where the CWA's capital mandate was binding, our results suggest that pairing grants with new regulation can increase the likelihood of full pass-through. This result may be generalizable to other intergovernmental transfer programs, such as grants for capital expenditure on highways and other infrastructure. This method of enacting policy is limited by the ability of donor governments to identify which receiving governments need fiscal assistance and estimate the costs of grant-funded programs. Both of these challenges to implementation hindered the CWA, leading to low grant pass-through.

Heterogeneity in pass-through across receiving municipalities can explain why our estimates are lower than previous estimates of CWA grant pass-through. In particular, Keiser and Shapiro (2019) leverages continuous variation in grant size in a difference-in-differences framework to show evidence of nearly full pass-through of CWA grants. This type of estimate is identified in part off of comparisons of municipalities that receive large grants to municipalities that receive small grants before and after grant receipt, but in order to identify a meaningful result using this type of comparison, we must assume that the effect of a grant of a given size is homogeneous across units (**?**). Because compliant and non-compliant municipalities respond to grant receipt differently, this assumption does not hold, and observations of compliant municipalities after grant receipt will not make valid counterfactuals for non-compliant municipalities (and vice-versa). We show that this causes estimates of CWA grant pass-through identified off of variation in grant size that do not account for heterogeneity in pass-through to over-estimate the effect of grant funding on targeted spending.

Similar problems can arise whenever there is heterogeneity in the relationship between grant amount and targeted spending across receiving municipalities. Evaluations of intergovernmental grant programs often estimate pass-through using difference-in-differences designs that leverage continuous variation in grant size, but pass-through is likely heterogeneous for many grant programs. For example, grant size is often positively correlated with receiving governments' preferences for public goods (Knight, 2002), so in many cases, the effect of a dollar of grant funding on targeted spending is increasing in grant size. This selection into grant size will bias pass-through estimates identified off of comparisons of units that receive large grants to units that receive small grants.

Finally, low pass-through in compliant municipalities motivates an examination of how these municipalities adjusted spending in response to grant receipt. CWA grants differ from many other intergovernmental transfers in that they support a function of local government that has a direct "fee-for-service" revenue stream where water and sewerage utilities collect charges from local residents. We show that compliant municipalities redistributed grant revenue to residents by reducing water and sewerage charges. This suggests that grant pass-through may be lower in settings where receiving governments control a channel through which they can redistribute money with relatively little friction and where costs of the public good are more salient to local residents than their benefits. Municipalities receiving federal grants for electrical utilities (USDA, 2019) or waste management (USDA, 2020), as well as those receiving modern wastewater treatment grants (Travis et al., 2004) may be able to use the same crowd-out mechanism.

Since annual historical spending data only exists for municipalities with populations over 75,000, we estimate pass-through for a relatively small number of municipalities. The municipalities in our sample are large, so our results represent the pass-through of $25.5 billion (2014$) of $153 billion total CWA grants. While we cannot directly estimate pass-through of CWA grants distributed to smaller municipalities, we provide evidence that municipal population is not correlated with pass-through for municipalities in our sample, which suggests that CWA grants outside of our sample may have a similar pass-through rate to the grants in our sample.

3.2 Background

3.2.1 The Clean Water Act

Congress originally passed The Clean Water Act in 1948 as the Federal Water Pollution Control Act. Facing pressure to enact policy to reduce surface water pollution after a series of high profile river fires, Congress significantly expanded the CWA in 1972. The strengthened CWA utilized both subsidies and new regulation to combat water pollution, distributing an estimated $153 billion (in 2014 dollars) in grants to fund upgrades to municipal wastewater treatment capital and imposing new capital standards on all municipal wastewater treatment plants in the United States.

The CWA funded wastewater treatment grants as follows. First, the EPA distributed grant money to states according to a congressionally mandated formula that was based on each state's wastewater treatment needs, as well as its total population and forecast population (Rubin, 1985). States then issued grants to municipalities according to priority lists that were based on the severity of nearby surface water pollution, the size of the population affected, the need for conservation of the affected waterway and that waterway's specific category of need.[1] States wrote their own priority lists, but these lists had to be approved by the EPA annually. The Clean Water Act also explicitly prohibited states from considering the geographic location of a receiving municipality, future population growth projections and any of a municipality's development needs that were not directly related to pollution abatement when writing priority lists (USEPA, 1980). Notably, federal and state governments did not explicitly allocate CWA grants according to the finances or spending preferences of receiving governments.

[1] We show that this did not lead to selection on treated potential outcomes in Appendix Section 3.7.1.1

Many CWA grants were specifically designated for upgrades to wastewater treatment technology. When the CWA came into effect, more than a quarter of all wastewater treatment facilities in the US were using relatively inexpensive primary treatment (USEPA, 2000), which filters out large detritus and improves the aesthetics of surface water, but discharges all but the heaviest organic material into rivers and lakes (USEPA, 1998). The CWA required all wastewater treatment facilities to upgrade to more effective secondary treatment technology, which removes about 85 percent of organic material before discharge, by 1977.[2] Additionally, many states required facilities to satisfy treatment technology requirements that were more stringent than the CWA's mandate (USEPA, 2000).[3]

While the benefits of upgrading a facility's treatment technology were well understood, large upfront capital costs often made these upgrades prohibitively expensive without federal support. Upgrading could cost as much as 30 percent of the initial cost of the facility (National Environmental Research Center, 1972), so in general, municipalities only invested in secondary treatment before the CWA came into effect because of either state level regulation or pressure from nearby communities to reduce the flow of harmful pollutants downstream (Jerch, 2018).

Many CWA grants were larger than the cost of upgrading treatment technology.[4] In an assessment of the wastewater treatment upgrades mandated by the CWA, the EPA estimated the cost of upgrading to secondary treatment as a function of the amount of wastewater flowing through a facility (USEPA, 1973). We use these estimates to calculate the cost of upgrading for each non-compliant facility that we have wastewater flow data. Figure 3.1 plots the total costs that each municipality faced against its total grant amount. This figure shows that, while grants were similar to costs for municipalities that received up to around $125 per capita in total grant aid, grants to municipalities that received more than this amount were substantially larger than the estimated cost of upgrading.[5]

[2]The EPA enforced the secondary treatment standard through the National Pollutant Discharge Elimination System (NPDES), the CWA's new permitting system, which required more than 65,000 industrial and municipal dischargers to obtain permits from the EPA or state governments. Permits required municipal treatment plants to employ secondary treatment and had to be renewed every five years. Violating the terms of a permit resulted in a compliance order or civil suit by the EPA, and violators could be fined up to $25,000 per day (Copeland, 2016).

[3]Our sample includes some facilities that were in compliance with the CWA's capital mandate but were not in compliance with more stringent state standards. These facilities still faced incentives to spend CWA grants on upgrades, so we classify them as non-compliant. In Appendix Section 3.7.4.2, we show that our results are robust to dropping municipalities bound by state level regulations that were more stringent than the CWA's.

[4]States distributed grant dollars above the cost of upgrading to fund non-mandated capital expenditures, such as expansions of facility capacity.

[5]Figures 3.8 and 3.9 apply quadratic and linear fits to this data.

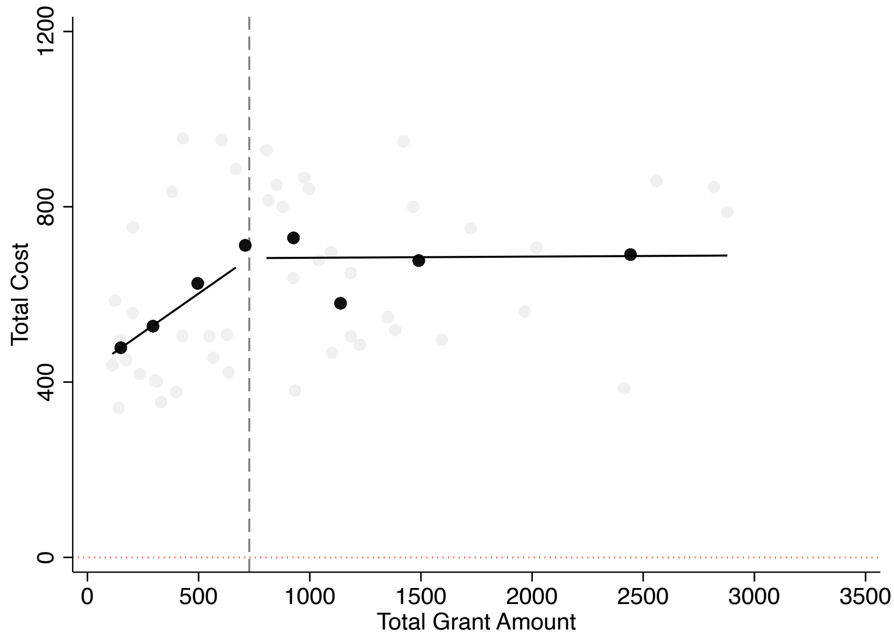


Figure 3.1: Costs Are Increasing in Grant Size Up to $125 and Constant Above $125

Notes: This figure plots estimates of non-compliant municipalities' total costs of upgrading to secondary treatment technology against the total grant dollars distributed to each municipality. The light gray dots are the true data and the black dots represent bins defined by deciles of total grant amount. Our cost estimates under-estimate costs for small grants, so we drop municipalities in the bottom decile of total grant aid. We divide the data at $125 per capita of total grant dollars, which is the point where costs are no longer increasing in grant size that we estimate in Table 3.3, and fit lines to the true data in each group.

Twelve percent of CWA grant dollars went to facilities that were already in compliance with both state and federal treatment technology standards. These grants were intended to fund non-mandated capital expenditures at wastewater facilities, such as capacity expansions.[6]

To illustrate variation in compliance and grant purpose, consider the example of two cities along the Willamette River in Oregon, historically one of the most polluted rivers in the Pacific Northwest. Eugene, OR is located at the top of the Willamette, while Portland lies at its base. To address municipal and industrial waste flowing downstream to Portland, the Oregon State Sanitary Authority (est. 1939) required Eugene and other municipalities discharging waste into the river to upgrade to secondary treatment in 1960. Because Eugene was bound by this state-level regulation, the city already met CWA treatment technology standards when they went into effect in 1972. Since Portland is located at the base of the Willamette, the city's discharge did not pollute other localities further downstream. Consequently, Portland was not subject to the 1960 treatment technology requirement and did not satisfy CWA treatment technology standards in 1972. Portland received CWA grant funding to upgrade its treatment facility to meet new federal standards in 1974. Although their facility already satisfied the CWA's capital mandate, Eugene received federal funding to expand facility capacity and upgrade one of its pump stations.

[6]The CWA allowed for grants to reimburse local governments that upgraded treatment technology prior to 1972 at a maximum 55% federal matching rate (USGAO, 1994), however, our grant data indicates that none of the complaint municipalities in our sample received grants for the purpose of reimbursement.

3.2.2 Grant Pass-Through

Despite its importance to policy design, researchers and policymakers alike remain unsure of when to expect grant programs to have higher or lower pass-through rates.[7] Economic theory predicts that local governments should treat inframarginal grant revenue (i.e. grant revenue smaller than existing local expenditure) as if it is an unconditional lump sum transfer (Oates, 1999), but empirical results suggest that this is not often the case, and that, even in the absence of binding constraints, inframarginal transfer revenue "sticks where it hits" (Henderson, 1968; Gramlich, 1969).[8] Known as the "flypaper effect," this phenomenon appears in a range of settings.[9] Proposed explanations for the flypaper effect include mis-perception by voters (Courant et al., 1979), mis-specification by researchers (Megdal, 1987), uncertainty (Vegh and Vuletin, 2015) and a focus by researchers on the short- rather than long-term effects of revenue changes on local spending (Helm and Stuhler, 2021), but a consensus has yet to emerge (Inman, 2008).

CWA grants did not place any explicit restrictions on local spending. For this reason, although CWA grants were allocated for specific projects, a municipality could redistribute these conditional grant funds toward other purposes if they were inframarginal to that municipality's budget. The median baseline sewerage expenditure in our sample ($20 per capita) was larger than the median yearly grant ($15 per capita), so a municipality could effectively disguise the continuation of existing spending as the use of grant funds while allocating grant dollars according to local preferences. Depending on local preferences, municipalities could spend inframarginal funds entirely on sewerage capital (full pass-through), entirely on other goods and services (full crowd-out), or a mix of sewerage capital and other goods and services.

CWA grants were matching grants. Grants covered 75% of wastewater facility project costs, while municipalities covered the remaining 25% themselves.[10] This local cost-sharing does not guarantee full pass-through, or even at least 25% pass-through, but rather increases the pre-grant spending necessary for grant funds to be considered inframarginal.[11] In particular, a 1:3 local-to-federal cost-sharing ratio implies that existing spending would need to be larger than $\frac{4}{3}$ times the grant amount for funds to be completely inframarginal.[12]

[7] In many cases, grant programs do not affect reduced form outcomes of interest unless they first cause an increase in spending. Gordon (2004) demonstrates this principle in the context of K-12 education by showing that previous findings of null effects of Title I transfers on student achievement can be explained by local governments reducing their own contribution to education spending in response to grant receipt.

[8] While this pattern often holds, estimates of grant pass-through are not consistent across settings (Hines and Thaler, 1995), and pass-through estimates can vary widely among similar grant programs. For example, Card and Payne (2002) studies the effects of school finance reforms between 1970 and 1992 on education spending and finds that a one dollar increase in state aid leads to a 60 cent increase in district education spending. Lutz (2010) examines a similar policy where state governments distributed unrestricted lump sum transfers for education to municipalities based on per-pupil property wealth and finds that these transfers are almost completely crowded out by changes in local spending.

[9] For example, researchers have documented the flypaper effect in intergovernmental grants for education (McGuire, 1978; Craig and Inman, 1986; Fisher and Papke, 2000; Card and Payne, 2002; Brunner et al., 2020), infrastructure (Gamkhar, 2000, 2003; Leduc and Wilson, 2017), welfare programs (Grossman and Roberts, 1989; Gamkhar and Oates, 1996) and police funding (Mello, 2019).

[10] In 1981, the federal share of costs was reduced to 55%.

[11] Cost-sharing yields two different possible interpretations of pass-through. We define full pass-through as $1 in grant funding leading to a $1 increase in sewerage capital spending. Alternatively, since municipalities are required to pay 25% of project costs, one could define full pass-through as $1 in grant funding leading to a $1.33 increase in sewerage capital spending (the actual $1 in grant aid plus the $0.33 in local obligations). Our estimates are based on the first definition, appealing to the notion of federal funding itself being spent as desired. We present pass-through estimates using the second definition in Appendix Section 3.7.1.2.

[12] The median baseline spending in our sample ($20 per capita) is approximately $\frac{4}{3}$ times the median yearly grant size ($15 per capita).

The interaction between the Clean Water Act's grant program and wastewater treatment technology mandate yields several predictions for our grant pass-through estimates. The Act's mandate on wastewater treatment technology has the potential to ensure full pass-through when grant amount is set equal to upgrade costs, but inaccurate grant sizing allows for crowd-out (when grants are too big) or imposes a fiscal burden on grant recipients (when grants are too small). The mandate imposes no constraint on how municipalities use inframarginal grant funds for compliant facilities.[13] Consequently, for cities already in compliance with the mandate, we predict pass-through near zero; for cities constrained by the mandate, we expect pass-through to decline in grant size relative to estimated costs.[14]

Finally, the contemporary political climate toward public utilities informs how municipalities may have redistributed grant funds that were not bound by the mandate. When the CWA came into effect, public utilities were facing pressure to reduce user fees and become more competitive (Daigger, 1998). Since local governments had the means to redistribute grants to residents with little friction by adjusting utility prices, CWA grants to compliant municipalities could crowd out sewerage capital spending already in place relatively easily.[15] Compliant municipalities could then use this crowded out money to offset water and sewerage utility costs usually funded through water bills.

3.2.2.1 Target Parameter

The term "grant pass-through" can describe a number of target parameters. We focus on the increase in targeted spending induced by the average grant dollar.[16] This parameter, which we refer to as the "total pass-through rate", is equivalent to the increase in *total* targeted spending in *all* receiving municipalities relative to the *total* grant dollars distributed to *all* receiving municipalities. While this average can mask substantial heterogeneity, it allows for an accounting of the effectiveness of an entire grant program. It also has an interpretation that is convenient for conducting benefit to cost analyses; dividing the benefit to cost ratio of all grant dollars by the total pass-through rate will return an estimate of the cost-effectiveness of grant dollars that receiving municipalities actually spent on the targeted category of expenditure. Our research design will aim to recover the total pass-through rate of CWA grants.

3.3 Data

3.3.1 Clean Water Act Data

We obtain data on every CWA grant from the EPA's Grant Information Control System. This data contains information on the year that the EPA distributed each grant, which municipality received the grant, the specific

[13] Compliant municipalities could still spend grant dollars on sewerage capital, but if they used this money to pay for upgrades that they already planned on making under their pre-grant budget constraint, CWA grants would not necessarily cause an increase in spending.

[14] Table 3.10 shows that compliant municipalities, non-compliant municipalities that received large grants and non-compliant municipalities that received small grants were observably similar before treatment. Table 3.11 shows that observable characteristics from 1971 do not predict grant size or compliance with the CWA's treatment technology mandate.

[15] The CWA instructed the EPA to issue guidelines to local governments on how to charge industrial and non-industrial users for waste treatment services, but did not place restrictions on adjusting rates.

[16] This is often the most policy-relevant parameter, but we may be interested in other parameters. For example, we may be interested in the increase in spending in the *average municipality* rather than the increase in expenditure induced by the *average dollar*. We estimate this "average pass-through rate" in Appendix Section 3.7.1.3.

wastewater treatment facility the grant was designated for and the amount distributed.[17]

Using a unique facility code, we merge our grant data to the 1972 Clean Watershed Needs Survey (CWNS), an assessment of the capital investment that publicly-owned wastewater treatment facilities required to come into compliance with the Clean Water Act. Importantly, the EPA conducted the 1972 CWNS before distributing any CWA grants, so the CWNS provides information on the treatment technology that each facility had in place before grant receipt.

The CWNS also includes information on the total amount of wastewater flowing through many facilities. We use these flow rates to calculate the costs that each non-compliant facility faced to upgrade its wastewater treatment technology using cost estimates from USEPA (1973), a separate EPA analysis of upgrade costs.[18]

3.3.2 Municipal Finance Data

We merge this linked facility level data to spending data from the municipality that operates each facility, dropping facilities operated by municipalities that we do not have spending data for. We refer to municipalities with only compliant facilities as compliant municipalities and those with only non-compliant facilities as non-compliant municipalities. Since we cannot observe exactly how local governments spend grants for specific treatment plants, we drop the 45 municipalities in our finance data with both compliant and non-compliant facilities from our sample.[19] Many municipalities receive multiple grants during our study period, so we define treatment as an absorbing state that begins when a municipality receives its first CWA grant.

Our data on municipal finances comes from the Census Bureau's Historical Database on Individual Government Finances. This database includes annual financial data starting in 1951 sourced from the *Compendium of City Government Finances* and *City Government Finances* surveys. Initially the Census Bureau surveyed the universe of municipalities with a population of at least 25,000. This limit rose to 50,000 in 1960 and 75,000 in 1987. As a result, we observe spending in 216 municipalities every year from 1951 until 1999.[20]

[17]Our analysis does not include grants distributed through predecessor programs similar to the CWA. See Appendix Section 3.7.4.2 for further discussion.

[18]We do not calculate cost estimates for the 23 non-compliant municipalities in our sample that operate facilities that are missing information on total wastewater flow.

[19]Some facilities in the CWNS are missing information on compliance with treatment technology standards, however, no facilities operated by municipalities in our spending data are missing this information.

[20]Starting in 1967, the Census also collected spending data from a larger sample of municipalities once every five years. To obtain accurate pass-through estimates, we need to measure spending in all post-treatment periods, so we cannot use this data to estimate pass-through. We re-estimate the crowd-out results from Section 3.5 using this data in Appendix 3.7.2.

Table 3.1: 1970 Summary Statistics by Facility Compliance

	(1) Compliant	(2) Non-Compliant	(3) Difference
Population	115,501	396,525	-281,023
	(73,126)	(954,559)	(204,352)
Total Revenue	253.13	294.21	-41.08
	(138.17)	(155.15)	(36.16)
Total IGR	52.84	64.39	-11.55
	(62.93)	(70.45)	(16.43)
Federal IGR	11.43	10.85	0.58
	(15.64)	(11.56)	(2.96)
Nonwater Utility Rev	21.18	18.15	3.03
	(44.79)	(45.12)	(10.72)
Property Tax Revenue	82.21	98.30	-16.09
	(72.75)	(79.08)	(18.53)
Water Utility Revenue	23.10	22.91	0.19
	(9.10)	(8.58)	(2.06)
Total Expenditure	277.26	308.16	-30.90
	(154.56)	(166.63)	(39.10)
Total IG Exp	1.93	4.53	-2.60
	(4.54)	(9.98)	(2.19)
Capital Outlays	79.35	63.88	15.47
	(57.33)	(37.65)	(10.02)
Highway Exp	21.16	21.95	-0.79
	(8.38)	(9.41)	(2.19)
Sewerage Capital Outlay	8.94	6.17	2.77
	(9.61)	(6.80)	(1.76)
Sewerage Non-capital Exp	5.50	5.04	0.46
	(3.13)	(3.11)	(0.74)
Nonwater Utiltiy Exp	22.46	18.82	3.64
	(49.25)	(46.86)	(11.26)
Water IG Exp	1.98	2.39	-0.40
	(1.91)	(2.59)	(0.59)
Water Non-capital Exp	13.34	12.85	0.49
	(4.74)	(5.90)	(1.36)
Water Capital Outlay	13.45	8.37	5.08
	(19.01)	(11.53)	(3.16)
Observations	22	90	112

Notes: This table presents summary statistics for municipalities with compliant and non-compliant facilities in 1970, two years prior to the CWA. All covariates aside from the facility compliance dummy and population represent per capita values.

Variables concerning municipal spending on water and sewerage utilities are available beginning in 1956. The data separates expenditure toward sewerage services into capital outlays and non-capital expenditure. We also observe revenue raised through a city's water and sewerage utility services (i.e. water bills). Because timing of grant receipt varies, we adjust both grant amount and all municipal spending outcomes to 1973 dollars.

Figure 3.2: Map of Municipalities

Notes: This figure plots the locations of the 112 municipalities in our sample. Non-compliant municipalities are shown as Xs and compliant municipalities are shown as points.

States had some discretion about where they distributed grants, so grant receipt might be correlated with trends in municipal spending. Rather than leveraging variation in grant *receipt*, we focus our analysis on the municipalities in our finance data that received at least one CWA grant and estimate pass-through by leveraging variation in grant *timing*.[21] Table 3.1 presents summary statistics for these 112 municipalities from 1970, two years before Congress passed the CWA, split based on whether or not the city's wastewater treatment facilities were in compliance with the CWA's capital mandate when the Act came into effect.[22] All variables besides population are in per capita terms. Compliant and non-compliant municipalities demonstrate no statistically distinguishable differences across any of the observable characteristics presented in Table 3.1, though this is largely due to statistical imprecision from the small sample size. Comparisons of means suggest that, on average, non-compliant municipalities in our sample had larger populations and greater total revenues and expenditures than compliant municipalities. Taken at face value, transfer revenue from other governments to non-compliant municipalities and more lucrative property taxes explain this difference. Figure 3.2 shows the location of each municipality in our sample.[23]

3.4 Grant Pass-Through

3.4.1 Pass-Through Methods

We want to select a design that recovers the total pass-through rate. To do this, we need an estimate of the average treatment effect on the treated (ATT), which describes the average increase in sewerage capital

[21]We obtain similar results if we use the municipalities in our finance data that never received a CWA grant as a control group in Appendix Section 3.7.1.4.

[22]We show that observable spending characteristics do not predict grant timing in Appendix Section 3.7.1.5.

[23]See Appendix 3.7.4 for more discussion of our data.

expenditure after grant receipt. Dividing the ATT by average grant amount returns the pass-through rate for the average CWA grant dollar in our sample.[24]

We follow Callaway and Sant'Anna (2020b) and leverage variation in the timing of CWA grant receipt to estimate "group-time average treatment effects". At least one municipality receives its first CWA grant in every year from 1972 to 1981, as well as in 1985 and 1988. This variation in treatment timing yields 12 "timing groups", which we index $g = 1972, ..., 1981, 1985, 1988$. We estimate the group-time average treatment effect for each group g in each time period t by comparing units in g to units that were not yet treated in time t. We then summarize the reduced form effect of grant receipt on sewerage capital spending by aggregating these group-time average treatment effects together.[25]

Under a parallel trends assumption, the group-time average treatment effect for group g at time t is

$$ATT(g,t) = E[C_t - C_{g-1} | G_g = 1] - E[C_t - C_{g-1} | D_t = 0]$$

where C_t is average sewerage capital spending at time t, G_g is a dummy variable that equals one for units in timing group g and D_t is a dummy that equals zero for units not-yet-treated at time t. We estimate each $ATT(g,t)$ with its sample analogue, $\widehat{ATT(g,t)}$.[26] This process yields many $\widehat{ATT(g,t)}$, most of which are identified off of relatively few observations, so instead of interpreting individual $\widehat{ATT(g,t)}$, we summarize the effect of treatment by aggregating the $\widehat{ATT(g,t)}$ together.

We aggregate group-time treatment effects into summary measures in several ways. First, we examine how the effect of CWA grant receipt on sewerage capital expenditure evolves over time with an event study specification. For each time relative to treatment e, we estimate the effect of treatment for units that have been treated for e periods with $\widehat{\theta}_D(e)$.

$$\widehat{\theta}_D(e) = \sum_{g=1972}^{1985} \sum_{t=1957}^{1987} 1\{t - g = e\} \widehat{ATT(g,t)} P(G = g | t - g = e) \tag{3.1}$$

[24] Two-way fixed effects (TWFE) estimators are one of the most common ways to leverage variation in treatment timing. This estimator regresses targeted spending on a treatment dummy and unit and time fixed effects. The resulting estimate is a weighted average of comparisons between (1) spending in newly treated municipalities relative to spending in municipalities that have not yet been treated and (2) spending in newly treated municipalities relative to spending in already-treated municipalities. While the first type of comparison only requires a parallel trends assumption to recover an estimate of the ATT, to use the second type of comparisons, we must assume that treatment effects are constant over time. If treatment effects are dynamic, already-treated units, which are still actively responding to treatment, will not make valid counterfactuals for what would have happened in the absence of treatment, so the second type of comparison does not return a meaningful result and we cannot interpret a TWFE coefficient as the ATT (Goodman-Bacon, 2021a). We have an a priori reason to expect the effect of grant receipt on spending to change over time; states distributed CWA grants to fund specific projects, so we would expect grant receipt to cause an increase in sewerage capital spending that only lasts until the project is completed. When spending returns to pre-treatment levels, this decrease is subtracted from a TWFE estimate, biasing the estimate upwards. Since this estimator does not return the ATT, we cannot recover the total pass-through rate from a TWFE estimate, which motivates us to use a different estimator.

[25] Since this estimator only relies on comparisons between newly treated units and not-yet-treated units, we do not need to place any restrictions on treated potential outcomes, and our estimates will not be biased by dynamic treatment effects.

[26] $\widehat{ATT(g,t)} = E_n[\widehat{C}_t - \widehat{C}_{g-1} | G_g = 1] - E_n[\widehat{C}_t - \widehat{C}_{g-1} | D_t = 0]$.

Taking an average of the $\widehat{\theta}_D(e)$ for e such that $e \geq 0$ provides a summary measure of the reduced form effect of grant receipt on sewerage capital spending.

Another way to summarize this effect is to take a weighted average of all of the $\widehat{ATT(g,t)}$ such that $g \leq t$ where the weights are based on timing group size. The formula for this weighted average is

$$\frac{1}{\kappa} \sum_{g=1972}^{1985} \sum_{t=1972}^{1987} 1\{g \leq t\} \widehat{ATT(g,t)} P(G=g) \tag{3.2}$$

where $\kappa = \sum_{g=1972}^{1985} \sum_{t=1972}^{1987} 1\{g \leq t\} P(G=g)$.

We can use these estimates to recover the total pass-through rate. Under a parallel trends assumption, our summary measures of the effect of grant receipt on sewerage capital spending represent estimates of the average increase in targeted spending after grant receipt. We can repeat our estimation process with grant amount as the dependent variable to estimate the associated first stage. Dividing the reduced form effect of grant receipt on sewerage capital spending by the first stage relationship between grant receipt and grant amount returns the average increase in sewerage capital spending induced by a dollar of grant funding, which is the total pass-through rate.

To identify a given $ATT(g,t)$, we must assume that, in absence of treatment, the average outcomes in timing group g would have followed parallel trends with not-yet-treated groups in period t. Since we are primarily interested in summary measures, we do not need to interpret individual $\widehat{ATT(g,t)}$ and we do not need this assumption to hold for all group-time pairs. Instead, we must assume that, on average, parallel trends holds for all groups g in all periods t such that $g \leq t$.

Callaway and Sant'Anna (2020a) constructs standard errors with a multiplier bootstrap procedure. Instead of re-sampling observations as in a pair-bootstrap, each multiplier bootstrap draw perturbs the influence function of the estimate (which measures the dependence of the estimate on each cluster in the sample). Our pass-through estimator is a function of two Callaway and Sant'Anna (2020a) objects (specifically, the reduced for divided by the first stage), and rather than derive the influence function for this estimator, we present pair-bootstrap standard errors clustered at the municipality level for our pass-through estimates.

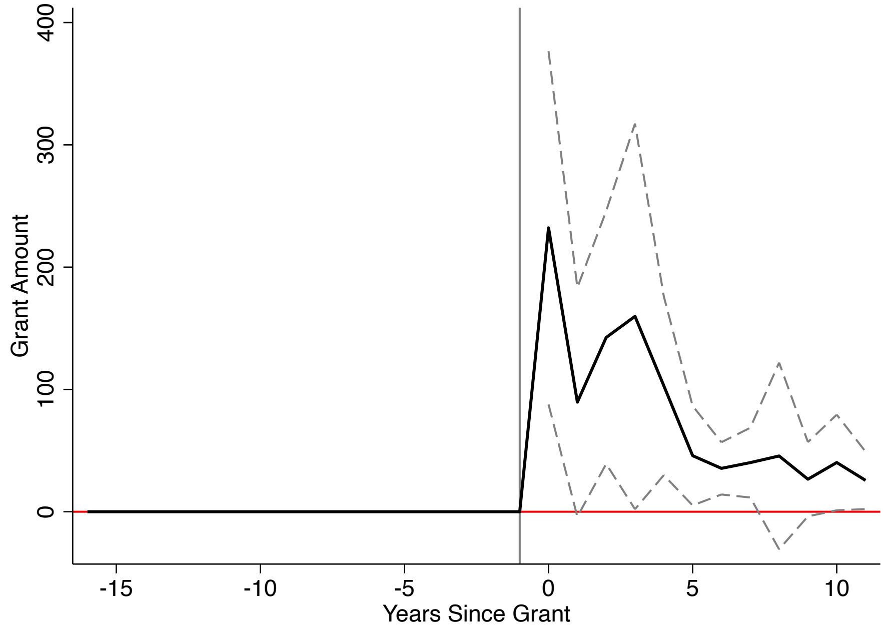


Figure 3.3: First Stage Relationship Between Grant Receipt and Grant Amount

Notes: This figure shows average grant amount for each year relative to treatment.

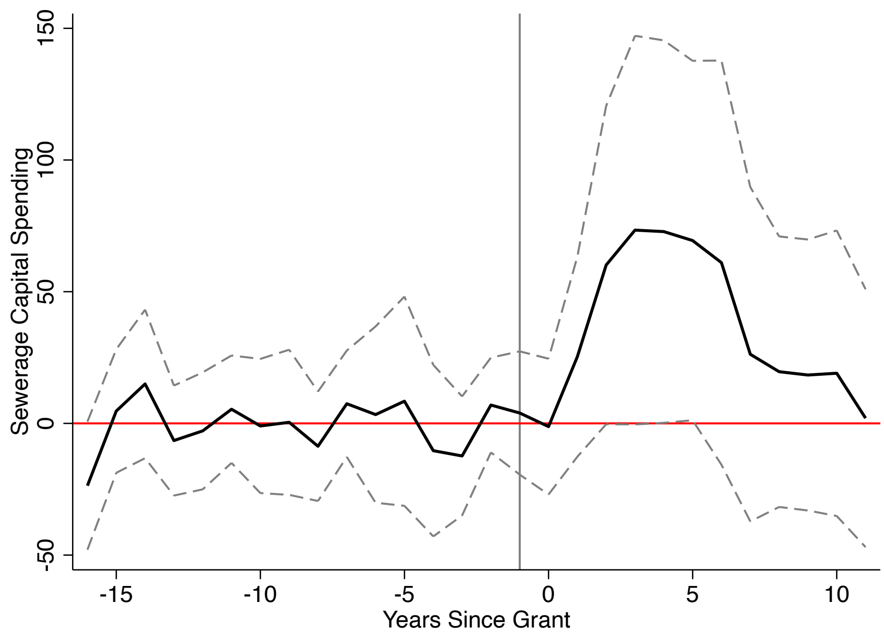


Figure 3.4: Sewerage Capital Spending Increased After Grant Receipt

Notes: This figure shows the reduced form relationship between grant receipt and sewerage capital spending by presenting the $\widehat{\theta}_D(e)$ for each year e relative to treatment, where $\widehat{\theta}_D(e) = \sum_{g=1972}^{1985} \sum_{t=1957}^{1987} 1\{t-g=e\}\widehat{ATT(g,t)}P(G=g|t-g=e)$. g indexes timing groups and t indexes years. $\widehat{ATT(g,t)}$ is the difference between changes in spending in units in group g in periods t and $g-1$ and changes in spending in units not yet treated at time t in periods t and $g-1$.

Table 3.2: CWA Grants had Low Pass-Through

	(1)	(2)	(3)
Panel A: Dynamic Aggregation	Grant Amount	Sewerage Capital	
Treat * Post	11.72***	5.628***	
	(1.31)	(1.746)	
Pass-through			0.480**
			(0.207)
p-value: Pass-through = 0			0.020
p-value: Pass-through = 1			0.012
Observations	3630	3630	3630
Panel B: Dynamic Aggregation ($e \leq 6$)			
Treat * Post	19.81***	8.844***	
	(2.768)	(2.485)	
Pass-through			0.446***
			(0.142)
p-value: Pass-through = 0			0.002
p-value: Pass-through = 1			0.000
Observations	3630	3630	3630
Panel C: Balanced Dynamic Aggregation			
Treat * Post	19.87***	8.914***	
	(3.011)	(2.302)	
Pass-through			0.448***
			(0.137)
p-value: Pass-through = 0			0.001
p-value: Pass-through = 1			0.000
Observations	3564	3564	3564
Panel D: Simple Aggregation			
Treat * Post	13.26***	6.115***	
	(1.526)	(1.980)	
Pass-through			0.461***
			(0.156)
p-value: Pass-through = 0			0.003
p-value: Pass-through = 1			0.001
Observations	3630	3630	3630

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the first stage relationship between grant receipt and grant amount, the reduced form effect of grant receipt on sewerage capital spending and the pass-through rate for all CWA grants. Panel A presents averages of the $\widehat{\theta}_D(e)$ for each time relative to treatment e for $e \geq 0$ calculated with $\widehat{\theta}_D(e) = \sum_{g=1972}^{1985} \sum_{t=1957}^{1987} 1\{t-g=e\}\widehat{ATT(g,t)}P(G=g|t-g=e)$. Panel B re-estimates the results in Panel A for e such that $0 \leq e \leq 6$. Panel C re-estimates the results in Panel B for municipalities that we observe for at least seven post-treatment periods. Panel D presents averages of group-time treatment effects based on group size calculated with the following equation $\frac{1}{\kappa} \sum_{g=1972}^{1985} \sum_{t=1972}^{1987} 1\{g \leq t\}\widehat{ATT(g,t)}P(G=g)$ where $\kappa = \sum_{g=1972}^{1985} \sum_{t=1972}^{1987} 1\{g \leq t\}P(G=g)$. Grant amount is the dependent variable in column 1, and sewerage capital spending is the dependent variable in column 2. Column 3 shows the implied pass-through rate from dividing the reduced form in column 2 by the first stage in column 1.

3.4.2 Full Sample Pass-Through Results

We begin with a flexible specification that explores how the effect of grants evolves over time. Figure 3.3 examines the first stage relationship between grant receipt and grant amount by presenting the $\widehat{\theta}_D(e)$ from equation 3.1 with grant amount as the dependent variable. Each $\widehat{\theta}_D(e)$ in Figure 3.3 is equal to average grant amount e years after treatment. Figure 3.4 then presents the associated reduced form estimates of the effect of grant receipt on sewerage capital spending. The null effects in the 16 years before grant receipt support a research design that leverages variation in grant timing by showing common trends in sewerage capital spending in treated and not-yet-treated municipalities before treatment, which suggests that parallel trends would have continued in absence of treatment. The estimates increase after the arrival of a municipality's first CWA grant and remain high for seven years after grant receipt. Sewerage capital spending decreases to near pre-treatment levels by eight years after grant receipt. The EPA estimates that upgrades paid for with CWA grants could take up to 10 years from initial grant receipt to project completion (USEPA, 2002), so the shapes of Figures 3.3 and 3.4 are consistent with grant funding and sewerage capital spending returning to pre-treatment levels as municipalities complete grant funded projects.[27]

Table 3.2 summarizes the results in Figures 3.3 and 3.4. First, Panel A presents averages of the $\widehat{\theta}_D(e)$ from Figures 3.3 and 3.4. Since the effect of grant receipt on both sewerage capital spending and grant amount return to near pre-treatment levels by seven years after treatment, Panel B re-calculates the results in Panel A for the fist seven post-treatment periods.

Since treatment timing varies, we do not observe all of the municipalities in our sample for the same number of post-treatment periods. For this reason, the composition of the groups that contribute to each event study coefficient in Figures 3.3 and 3.4 changes when event time changes. This causes our dynamic aggregations in Panels A and B of Table 3.2 to weight up observations of municipalities that are treated earlier, which can produce misleading results if the effect of grant receipt varies across timing groups. We address this in Panel C by re-calculating the results in Panel B on the sub-sample of municipalities that we observe for at least seven post-treatment periods. The estimates are similar to those in Panel B.[28]

Finally, Panel D of Table 3.2 presents a weighted average of all post-treatment group-time-treatment effects calculated with equation 3.2.

Dividing the reduced form estimates in Table 3.2 by their respective first stage estimates suggests that CWA grants have a total pass-through rate between 0.446 and 0.480. Across aggregations, the 95% confidence intervals of the first stage and reduced form estimates do not overlap and we can reject a test of the hypothesis that our pass-through estimate equals one, so we reject full grant pass-through for the full sample.

[27] Table 3.12 presents the results in Figures 3.3 and 3.4 in tabular form.
[28] Figure 3.10 presents the associated event study.

3.4.3 Evaluating CWA Investments

We can use estimates of the total pass-through rate for the full sample to find the benefit to cost ratio of the CWA grant dollars that municipalities spent on sewerage capital (i.e., the benefit to cost ratio of the wastewater capital upgrades themselves). Using increased housing prices to quantify the benefits of CWA grants, Keiser and Shapiro (2019) estimates a benefit to cost ratio of 0.26. Flynn and Marcus (2021) finds that this ratio may be as high as 0.45 after incorporating improvements to infant health.[29] Dividing this benefit to cost ratio by the total pass-through rate yields the benefit to cost ratio of CWA grant dollars that municipalities spent on sewerage capital. Our preferred specification in Panel B of Table 3.2 shows that, in total, receiving governments in our sample spent 44.6% of CWA grant dollars on sewerage capital upgrades. Dividing these results implies that the CWA grant dollars that municipalities spent on sewerage capital have a benefit/cost ratio of 1.01.[30] This suggests that, given previously quantified benefits of the CWA, accounting for low pass-through can lead to favorable benefit to cost ratios of CWA-funded wastewater capital upgrades.[31]

The standard error on our preferred specification suggests a 95% confidence interval of 0.165 to 0.727. Dividing previously estimated benefits of CWA grants by the upper bound lets us reject a benefit/cost ratio less than 0.62.

We examine which municipalities are driving this low pass-through in the next section.

3.4.4 Heterogeneity in Pass-Through

We explore heterogeneity in pass-through by mandated costs. Compliant municipalities face no mandated costs, and mandated costs vary within our sample of non-compliant municipalities. Since, in theory, setting grant amount equal to mandated costs increases the likelihood of full pass-through, we estimate pass-through separately for compliant municipalities, non-compliant municipalities where grant amount is approximately equal to mandated costs, and non-compliant municipalities where grant amount is greater than mandated costs.

Splitting the sample along compliance is straightforward, but it is not immediately obvious how to divide non-compliant municipalities based on mandated costs. We first need to identify which grants are approximately equal to mandated costs and which grants are larger than mandated costs. Once we define these groups, we can separately estimate pass-through for (1) non-compliant municipalities whose grant funding is completely bound by the capital mandate and (2) non-compliant municipalities that receive grants that include funding that is not bound by the capital mandate. Figure 3.1 shows that, while grants are similar to

[29]This assumes that hedonic estimates do not capture any health benefits. It is unlikely that this historical population fully understood the relationship between surface water quality and the health of infants in utero, so this assumption is plausible.

[30]The other results in Table 3.2 suggest that the CWA grant dollars that municipalities spent on sewerage capital had a benefit/cost ratio of between 0.937 and 1.01. This benefit/cost ratio is even higher if we use the estimates from our stacked difference-in-difference in Table 3.24, which suggest a benefit/cost ratio of 1.32.

[31]We discuss potential shortcomings of this back-of-the-envelope calculation in Appendix Section 3.7.1.7.

costs for municipalities that receive relatively small grants, grants to municipalities that receive large grants are substantially larger than the estimated cost of upgrading. For this reason, we want to divide our sample of non-compliant municipalities into those that received large and small grants.

We divide our sample at the point on the distribution of grant amount where costs are no longer increasing in grant size.[32] To estimate this point, we keep one observation from each non-compliant municipality that we have cost data for and estimate the following equation with non-linear least squares

$$\begin{aligned} cost_i = &(a_1 + b_1 * TotalGrant_i) * 1\{TotalGrant_i < split\} \\ &+ (a_1 + b_1 * split + b_2 * (TotalGrant_i - split)) * 1\{TotalGrant_i >= split\} \end{aligned} \tag{3.3}$$

where $cost_i$ is our estimate of municipality i's total per capita cost of upgrading and $TotalGrant_i$ is municipality i's total per capita grant amount. Column 1 of Table 3.3 presents our estimate of $split$. This result shows that costs are increasing in grant size up to \$125 of total per capita in grants and flatten out for grants above \$125. We estimate pass-through separately for non-compliant municipalities above and below this threshold.[33]

Table 3.3: Non-Linear Least Squares Estimates of Break in Cost

	(1) Total Cost	(2) Total Cost
Split	125.2**	138.2
	(60.49)	(106.8)
Observations	41	58

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table presents $split$ from estimating $cost_i = (a_1 + b_1 * grant_i) * 1\{grant_i < split\} + (a_1 + b_1 * split + b_2 * (grant_i - split)) * 1\{grant_i >= split\}$ on a sample of non-compliant municipalities with non-linear least squares. Column 1 presents results from estimating this equation with total cost calculated from estimates in USEPA (1973) as the dependent variable. Column 2 re-estimates the result in column 1 without dropping municipalities in the bottom decile of total grant amount.

3.4.4.1 Semi-Parametric Methods

Rather than imposing parametric restrictions on the relationship between grant amount and targeted spending, we explore heterogeneity in pass-through using an approximation of the dose-response function relating grant amount and sewerage capital spending. As shown in Figure 3.1, there is substantial variation in grant amount,

[32] This is the point where the quadratic fit in Figure 3.8 flattens out.

[33] The EPA did not include facilities with a total flow rate of less than 5 million gallons per day in their sample when estimating the cost functions that we use to calculate costs. These cost functions over-estimate costs for smaller facilities (USEPA, 1973). Most municipalities in the bottom decile of grant aid had facilities of this size, so we drop these municipalities when we estimate the results in column 1 of Table 3.3. We re-estimate equation 3.3 on the full sample in column 2, which yields an estimated cutoff of \$138 per capita. Table 3.13 shows that our pass-through estimates are similar if we use either of the cutoffs in Table 3.3.

and plotting out each municipality's total grant amount will give us the x-axis of this dose-response function. We can then find the y-axis by estimating each municipality's response to grant receipt.

To do this, we follow Deshpande and Li (2019) and re-organize our data into "stacks". Each stack S is defined by a single treated municipality, called municipality s, which is labeled as treated in that stack. We then add municipalities that received grants at least seven years after the treated municipality to the stack, which are labeled as controls. We can estimate the reduced form effect of grant receipt on sewerage capital spending for municipality s by comparing spending between municipality s and the control municipalities in stack S before and after the year in which municipality s becomes treated.[34] Estimating equation 3.4 on observations in stack S returns the result from this comparison.

$$C_{it} = \alpha_0 + \beta_s^{2X2} post_t * treat_i + \delta post_t + \omega treat_i + \varepsilon_{it} \tag{3.4}$$

Municipalities are indexed by i, stacks by s, and years by t. C_{it} is per capita sewerage capital spending, and $treat_i$ is a dummy variable that equals one for municipality s (the treated unit). Plotting β_s^{2X2} against the total amount of grant dollars distributed to municipality s for all stacks yields a semi-parametric approximation of the dose-response relationship between grant amount and sewerage capital spending.[35]

We summarize these figures by appending groups of stacks into one dataset and estimating a stacked difference-in-difference. Equation 3.5 estimates first stage relationship between grant receipt and grant amount, denoted g_{it},

$$g_{it} = \alpha_0 + \beta_{fs}^{stacked} D_{it} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its} \tag{3.5}$$

and equation 3.6 estimates the reduced form effect of grant receipt on sewerage capital spending,

$$C_{it} = \alpha_0 + \beta_{rf}^{stacked} D_{it} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its} \tag{3.6}$$

where D_{it} is a dummy variable that equals one after grant receipt, and α_{is} and α_{ts} are stack-by-municipality

[34] By estimating the effect of grant receipt on *individual units* across *all post treatment periods*, a stacked estimator allows us to semi-parametrically examine the relationship between grant amount and sewerage capital spending before constructing summary measures of pass-through. Since timing groups contain both compliant and non-compliant municipalities, as well as municipalities that received different sized grants, we cannot repeat this process with Callaway and Sant'Anna (2020a) without making a priori assumptions about the relationship between grant amount and sewerage capital spending, as well as the relationship between compliance and sewerage capital spending.

[35] To identify the β_s^{2X2}, we must assume that sewerage capital spending in municipality s and the control municipalities in stack S would have followed parallel trends in absence of treatment. This is stronger than the assumption we need to identify $\beta_{rf}^{stacked}$ in equation 3.6, which only requires us to assume that parallel trends holds on average (we derive this identifying assumption in Appendix Section 3.7.3.4). For this reason, we should interpret individual β_s^{2X2} with caution. Note that, under parallel trends, each β_s^{2X2} represents the effect of going from zero grant dollars to the grant dollars distributed to municipality *s on municipality s*. This does not tell us anything about the effect of going from municipality $s-1$'s grant amount to municipality s's grant amount, or the effect of grant receipt on any other municipality.

and stack-by-year fixed effects. In tandem, theses fixed effects force identification to come from within-stack variation, ensuring that our estimates will not reflect any comparisons between newly treated municipalities relative and already-treated municipalities. Since it is only identified off of within-stack comparisons, $\beta_{rf}^{stacked}$ is an unweighted average of all the β_s^{2X2}. We present pair-bootstrap standard errors clustered at the municipality level.[36]

3.4.4.2 Pass-Through Results for Sub-Groups of CWA Grants

Figure 3.5 examines the relationship between grant amount and sewerage capital spending in non-compliant municipalities. This figure plots out the β_s^{2X2} from estimating equation 3.4 on units in each non-compliant stack against municipality s's total grant amount.[37] Sewerage capital spending is increasing in grant amount for municipalities that received grants totaling up to \$125 per capita, but this relationship is flat for municipalities receiving more than \$125 per capita. Importantly, Figure 3.5 shows a relationship between the β_s^{2X2} and total grant amount that is very similar to the relationship between cost and total grant amount shown in Figure 3.1.

Figure 3.6 repeats this process for compliant municipalities. This figure shows that grants to compliant municipalities had low pass-through regardless of size.

Table 3.4 summarizes Figures 3.5 and 3.6 by estimating equations 3.5 and 3.6 on sub-samples of stacks. Panel A presents pass-through estimates for non-compliant municipalities that received grants totaling up to \$125 per capita. The pass-through estimate in column 3 suggests that CWA grants led to a dollar-for-dollar increase in sewerage capital expenditure in these municipalities. Panel B presents estimates for all other stacks defined by non-compliant municipalities, which have much lower pass-through. Panel C presents estimates for all stacks defined by compliant municipalities, which also have low pass-through.

Interpreting the results in Table 3.4 as the total pass-through rate suggests that non-compliant municipalities in our sample that received grants totalling less than \$125 dollars per capita increased sewerage capital spending by 104.9% of all grant dollars distributed to them, while all other non-compliant municipalities in our sample increased sewerage capital spending by 26.7% of all grant dollars distributed to them. Compliant municipalities in our sample increased sewerage capital spending by 3.37% of grant dollars distributed to them.[38]

[36]To construct standard errors, we take a random sample (with replacement) of municipalities in the unstacked data, re-form stacks, then perform our estimates. We repeat this process 1000 times and use the results to calculate bootstrap standard errors. See Appendix Section 3.7.3.4 for further discussion of stack construction and inference with a stacked difference-in-differences estimator.

[37]Because they reflect averages across seven years, we multiply the β_s^{2X2} in Figures 3.5 and 3.6 by seven. This makes these figures more comparable in scale to Figure 3.1 and with one another.

[38]Table 3.14 shows that pass-through was low for compliant municipalities that received large and small grants alike, which suggests that the heterogeneity in pass-through by grant size documented in Panels A and B of Table 3.4 is driven by differences in costs rather than other confounding factors associated with grant size that are similar across compliant and non-compliant municipalities.

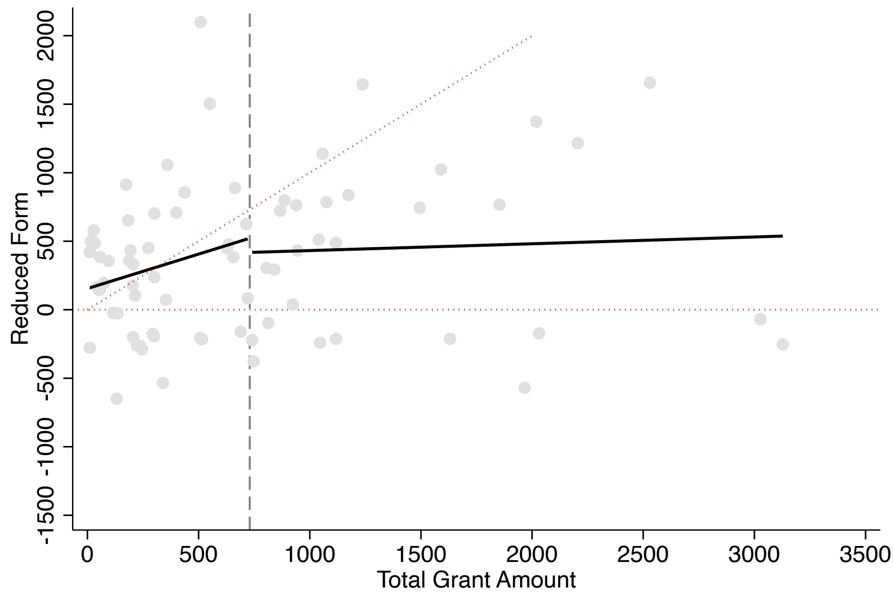


Figure 3.5: 2X2 DD Coefficients vs Grant Amount for Non-Compliant Municipalities

Notes: This figure plots the β_s^{2X2} from estimating $C_{it} = \alpha_0 + \beta_s^{2X2} post_t * treat_i + \delta post_t + \omega treat_i + \varepsilon_{it}$ on observations in each stack s (for stacks defined by non-compliant municipalities). Sewerage capital spending is the dependent variable. We plot each stack's β_s^{2X2} against the total grant dollars distributed to the treated municipality in that stack. We divide the stacks at \$125 per capita of total grant dollars and fit lines to the β_s^{2X2} in each group. We multiply the β_s^{2X2} by seven to reflect that each β_s^{2X2} represents an average across seven post-treatment years.

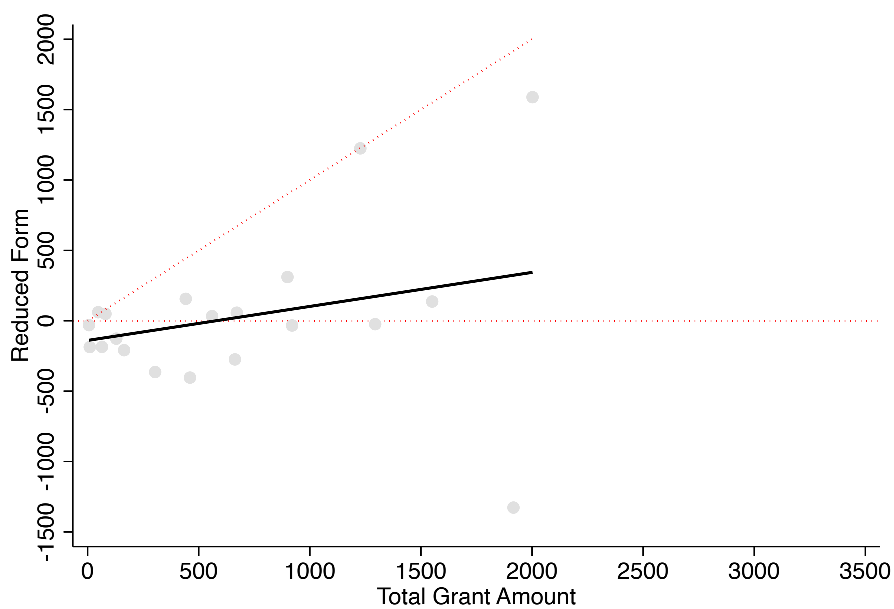


Figure 3.6: 2X2 DD Coefficients vs Grant Amount for Compliant Municipalities

Notes: This figure re-estimates the results in Figure 3.5 for stacks defined by compliant municipalities.

Table 3.4: Pass-Through is Heterogeneous in Grant Size and Compliance

	(1)	(2)	(3)
Panel A: Non-Compliant and < \$125	Grant Amount	Sewerage Capital	
Treat * Post	6.935***	7.279	
	(0.729)	(7.653)	
Pass-through			1.049
			(1.138)
p-value: Pass-through = 0			0.357
p-value: Pass-through = 1			0.965
Observations	2162	2162	2162
Panel B: Non-Compliant and >= \$125			
Treat * Post	46.01***	11.81	
	(6.306)	(7.130)	
Pass-through			0.267
			(0.156)
p-value: Pass-through = 0			0.087
p-value: Pass-through = 1			0.000
Observations	1518	1518	1518
Panel C: Compliant			
Treat * Post	16.43***	0.553	
	(3.905)	(4.655)	
Pass-through			0.0337
			(0.345)
p-value: Pass-through = 0			0.922
p-value: Pass-through = 1			0.009
Observations	920	920	920

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of $Y_{it} = \alpha_0 + \beta D_{it} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its}$. Municipalities are indexed by i, stacks by s, and years by t. D_{it} is a dummy variable that equals one after grant receipt, and α_{is} and α_{ts} are stack-by-municipality and stack-by-year fixed effects. Grant amount is the dependent variable in column 1, and sewerage capital spending is the dependent variable in columns 2 and 3. Column 3 uses D_{it} as a instrument for grant amount, which returns a coefficient equivalent to dividing the reduced form in column 2 by the first stage in column 1. Panel A presents estimates for stacks defined by non-compliant municipalities who received grants totalling less than \$125, Panel B for all other non-compliant stacks, and Panel C for all stacks defined by compliant municipalities. We present p values from testing for a pass-through rate of one at the bottom of each panel.

Based on the above results, we reject full pass-through for compliant municipalities and non-compliant municipalities receiving more than \$125 per capita in grant aid, but cannot reject full pass-through for non-compliant municipalities receiving grants totaling less than \$125 per capita, as expected.[39] While we cannot reject complete crowd-out in non-compliant municipalities that received grants totaling up to \$125 per capita or equality of the pass-through estimates in column 3 Table 3.4, these results, along with the shapes of Figures 3.5 and 3.6, are consistent with municipalities only increasing sewerage capital spending up to the point where

[39]We obtain similar results when we re-estimate the results in Table 3.4 with Callaway and Sant'Anna (2020a). We present these results in Appendix Section 3.7.1.6, and discuss our choice of estimators in Appendix 3.7.3.

they are in compliance with the CWA's capital mandate.[40]

3.4.5 Does Variation in Grant Size Recover Total Pass-Through?

Researchers often estimate pass-through by leveraging variation in grant size with a dose-response two-way fixed effects estimator. In this section, we follow **?** to show that, when the effect of a grant of a given size is not constant across receiving municipalities, or when the effect of grant receipt changes over time, this type of estimator may not return the total pass-through rate.

Consider, as an example, the pass-through estimates in Keiser and Shapiro (2019), which suggest that municipalities that received any CWA grant spent \$0.94 of every federal dollar on sewerage capital by estimating

$$C_{it} = \alpha_0 + \beta_{dr}^{TWFE} g_{it} + \alpha_i + \alpha_t + \varepsilon_{it} \tag{3.7}$$

where C_{it} is cumulative sewerage capital spending from 1970 to 2002 and g_{it} is cumulative grant amount. Similar to a binary TWFE estimate identified off of variation in treatment timing, β_{dr}^{TWFE} is a weighted average of (1) comparisons of newly treated municipalities relative to municipalities that have not yet been treated, (2) comparisons of newly treated municipalities relative to already-treated municipalities, and (3) comparisons of municipalities that received large grants relative to municipalities that received smaller grants.

To interpret comparisons of newly treated municipalities relative to already-treated municipalities as the ATT, we need to make additional assumptions along with parallel trends. When treatment effects are dynamic, already-treated units, which are still actively responding to treatment, will not make valid counterfactuals for what would have happened in the absence of treatment. For this reason, we must assume that treatment effects do not change over time to use this type of variation (Goodman-Bacon, 2021a).

Identifying a meaningful effect using comparisons of municipalities that received large grants relative to municipalities that received smaller grants requires additional assumptions as well. In particular, we must assume that a grant of a given size will have the same effect on sewerage capital spending in all receiving municipalities in our sample (**?**). If this assumption does not hold, municipalities that received small grants will not be valid counterfactuals for what would have happened if a municipality that received a large grant had received a small grant instead, and this type of comparison can either over- or under-estimate the total pass-through rate.

The results in Sections 3.4.2 and 3.4.4.2 suggest that the effect of CWA grants on sewerage capital spending is not homogeneous across time relative to treatment, or across receiving municipalities. As shown in

[40] We show that differences in pass-through between sub-groups are not driven by differences in baseline sewerage capital spending in Appendix Section 3.7.1.8.

Figure 3.4, treatment effects change over time as grant funding runs out. Additionally, Figure 3.5 suggests that the effect of grants to non-compliant municipalities is increasing in grant size (up to the point where the capital mandate is satisfied), while Figure 3.6 suggests that grants to compliant municipalities had little to no effect on sewerage capital spending regardless of grant size. For this reason, a grant of a given size would likely affect spending differently in a compliant municipality than in a non-compliant municipality, which violates the homogeneity assumption.[41]

The reliability of an estimate of β_{dr}^{TWFE} from equation 3.7 depends on how much of the estimate is identified off of comparisons of early treated units to not-yet-treated units. Table 3.5 decomposes β_{dr}^{TWFE} using Goodman-Bacon (2021b), which shows how much of β_{dr}^{TWFE} is identified off of comparisons of newly treated units to not-yet-treated units, comparisons of newly treated units to already treated units, and comparisons of units that received large grants to units that received smaller grants.[42]

Table 3.5: Decomposition of Keiser and Shapiro (2019) Pass-Through Estimates

	(1) Weight	(2) Average Estimate
Newly Treated vs Not-Yet-Treated	0.017	0.428
Newly Treated vs Already-Treated	0.553	1.002
Large Grant vs Small Grant	0.430	0.835

Notes: This table presents the average difference-in-difference estimates and their weights from decomposing β_{dr}^{TWFE} from $C_{it} = \alpha_0 + \beta_{dr}^{TWFE} + \alpha_i + \alpha_t + \varepsilon_{it}$ using Goodman-Bacon (2021b). C_{it} is cumulative sewerage capital spending, g_{it} is cumulative grant amount, and α_i and α_t are municipality and year fixed effects.
Source: Keiser and Shapiro (2019)

Row 1 of Table 3.5 describes the part of β_{dr}^{TWFE} identified off of comparisons of newly treated units to not-yet-treated municipalities. This variation suggests a pass-through rate of 0.428, which is similar to the pass-through estimates in Table 3.2.[43] This represents an estimate of the total pass-through rate, but less than two percent of β_{dr}^{TWFE} comes from this type of comparison.

Row 2 of Table 3.5 shows that more than half of β_{dr}^{TWFE} comes from comparisons of late treated municipalities to already-treated municipalities. Consistent with bias from treatment effects that decrease over time, comparisons of newly treated municipalities to already-treated municipalities yield a result that is more than

[41] Estimating β_{dr}^{TWFE} on compliant and non-compliant municipalities separately would solve the problem introduced by heterogeneity in treatment effects by compliance, but would not address problems caused by dynamic treatment effects. For this reason, estimating β_{dr}^{TWFE} on sub-groups of municipalities will not necessarily return the total pass-through rate for these sub-groups. See Appendix Section 3.7.3.2 for further discussion of the assumptions required to use this type of variation.

[42] This program is the analogue of Goodman-Bacon et al. (2019) for difference-in-difference estimates identified off of continuous variation.

[43] The part of the pass-through estimate in Keiser and Shapiro (2019) that is identified off of comparisons of newly-treated units to not-yet-treated is identified off of the same type of variation as the pass-through estimates in Table 3.2. The similarity of these results suggests that, even though we make several sample restrictions that Keiser and Shapiro (2019) do not (we drop both municipalities that are missing information on compliance and municipalities in our finance data with both compliant and non-compliant facilities), the differences between our estimates and those in Keiser and Shapiro (2019) come from differences in the source of variation we use rather than differences between our samples.

twice as large as the result from comparing newly treated municipalities to not-yet-treated municipalities.

Finally, Row 3 of Table 3.5 shows that a large portion of β_{dr}^{TWFE} comes from comparisons of municipalities that received large grants to those that received smaller grants. Many of these comparisons are between compliant and non-compliant municipalities, which do not represent causal estimates.

The decomposition in Table 3.5 shows that, when the effect of a grant of a given size is heterogeneous across receiving municipalities, estimating pass-through with a dose-response TWFE estimator will produce an estimate that may not equal the total pass-through rate.

Table 3.6: Pass-Through Results are Robust to Population Weighting

	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	14.55***	5.628***	
	(3.45)	(1.69)	
Pass-through			0.387**
			(0.162)
p-value: Pass-through = 0			0.017
p-value: Pass-through = 1			.000
Observations	3630	3630	3630

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results from Panel B of Table 3.2 weighting by population.

Table 3.7: Full Sample of Grants

Panel A: All Grants	(1) All	(2) Non-compliant	(3) Compliant
Total Grant Amount	2,741,595	2,454,547	2,260,387
Total Cost		932,683	
Observations	10,202	9027	940
Panel B: Grants with Spending Data			
Total Grant Amount	46,883,636	54,197,768	17,627,110
Total Cost		28,815,210	
Observations	112	90	22

Notes: Column 1 presents average total costs and average total grant amount for all municipalities that received CWA grants, and column 2 presents average total costs and average total grant amount the municipalities in our sample.

3.4.6 Can We Generalize These Results to All CWA Grants?

Our results represent the pass-through of \$25.5 billion (2014\$) of the \$153 billion in grants funded by the CWA, and in this section, we discuss how generalizable our results are to CWA grants outside of our sample.

The Census Bureau collects annual finance data from the universe of municipalities with populations over 75,000, so our sample only includes grants distributed to relatively large cities. If population is correlated with preferences for sewerage capital spending, grants to municipalities outside of our sample might have a different pass-trough rate than grants to municipalities within our sample. We adjust for heterogeneity in pass-through by population in Table 3.6 by weighting our pass-through results by population. This does not change our results in a meaningful way, which suggests that, within our sample, treatment effects are not correlated with population.[44] Since the Census data that we use is the only comprehensive source of state and local government finance data collected on a nationwide scale (US Census Bureau, 2021), we cannot test if this homogeneity holds *outside of our sample*, but the similar pass-through rates across municipality size *within our sample* suggests that CWA grant pass-through is not strongly correlated with population.

Since pass-through is correlated with mandated costs, our pass-through estimates might not be generalizable to all CWA grants if municipalities outside of our sample faced different mandated costs than municipalities in our sample. We compare municipalities in our sample to municipalities outside of our sample in Table 3.7. Municipalities outside of our sample received smaller grants than those in our sample, but they also faced lower mandated costs, so the ratio of grant size to mandated costs in municipalities within and outside of our sample is relatively similar. It is not unreasonable to expect non-compliant municipalities outside of our sample to increase sewerage capital spending up until the point where they are in compliance with the CWA's capital mandate, then reduce their own contribution to sewerage capital spending. If this is the case, then the similar relationship between grant size and mandated costs in non-compliant municipalities within and outside of our sample suggests that non-compliant municipalities outside of our sample should have a similar pass-through rate to non-compliant municipalities within our sample. Similarly, since compliant municipalities outside of our sample did not face any mandated costs, we would expect them to have pass-through close to zero.[45]

For the full sample of CWA grants, receiving governments include cities, towns, and sewage districts (Keiser and Shapiro, 2019), but our sample of municipal spending data only includes cities. Focusing on cities should not affect the generalizability of our results to CWA grants outside of our sample. Towns could use the same crowd-out mechanism as cities. In cases where grants are distributed to special water utility districts that handle wastewater treatment for multiple cities, no single city has authority to divert grant funds received by the utility district to other purposes, and the only diversion possible is for the utility district to

[44]We test for this more directly in the Appendix. Table 3.15 re-estimates our pass-through results on sub-samples of municipalities divided by tercile of population. While less precise, these estimates are relatively similar to our full sample estimates.

[45]In total, twelve percent of CWA grant dollars were distributed to facilities that were already in compliance with both state and federal treatment technology standards. Assuming that our estimates of the total pass-through rate for compliant and non-compliant municipalities are accurate, we can take a weighted average of pass-through estimates for compliant and non-compliant municipalities to approximate pass-through for all CWA grants. We present the results of this exercise in Table 3.16. This suggests a pass-through rate of 0.353, and we can reject full grant pass-through.

pass on grant funds to customers via lower fees, which is consistent with the crowd-out mechanism that we explore in the next section.

3.5 How did Municipalities Spend Crowded-Out Funds?

3.5.1 Redistribution Methods

Evidence of low pass-through in compliant municipalities motivates an examination of how these municipalities adjusted spending in response to grant receipt. Both Callaway and Sant'Anna (2020a) and our stacked difference-in-differences estimator rely on comparisons of newly treated units to not-yet-treated units, but since we observe a relatively small number of compliant municipalities, we leverage a different source of variation to estimate the effect of CWA grants on local spending in compliant municipalities. Since pass-through is generally lower in compliant municipalities than in non-compliant municipalities, we can estimate the effect of grant receipt on spending in compliant municipalities by comparing outcomes between compliant municipalities and non-compliant municipalities before and after grant receipt with equation 3.8.

$$R_{it} = \alpha_0 + \theta grant_{it} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it} \tag{3.8}$$

Our dependent variable of interest, R_{it}, is water utility revenue, which we expect to decrease as compliant municipalities redistribute money to residents by lowering utility bills. $grant_{it}$ is a dummy variable that equals one after a municipality receives a grant, $compliant_i$ is a dummy equaling one for compliant municipalities and α_i is a municipality fixed effect. g indexes timing groups, which are defined by the year in which a municipality receives its first CWA grant and α_{gt} is a timing-group-by-year fixed effect.

Including α_{gt} in equation 3.8 lets us estimate the effect of grants to compliant municipalities on spending without using any variation in treatment timing.[46] The result from equation 3.8 is equivalent to estimating the difference-in-difference in equation 3.9 on all observations in timing group g, which compares water revenues between compliant and non-compliant municipalities that receive their first grant in year g in periods before and after g, then averaging together the θ_g^{2X2} from all timing groups.

$$R_{it} = \alpha_0 + \theta_g^{2X2} 1\{t \geq g\} * compliant_i + \delta 1\{t \geq g\} + \omega compliant_i + \varepsilon_{it} \tag{3.9}$$

[46] As with our grant pass-through results, comparisons of early treated units to late treated units will be unbiased under parallel trends, but comparisons of late treated units to early treated units may be wrong signed. While there is nothing wrong with the first type of comparison, it is likely that variation in compliance will identify a different treatment effect than timing variation, so combining these sources of variation in one equation will reduce the precision of our estimates. Instead, we leverage these two sources of variation with different designs. We present results from a complementary design that identifies off of variation in treatment timing in Appendix Section 3.7.2.1.

We first examine the relationship between compliance and water revenue with an event study that compares water revenue in compliant municipalities to water revenue in non-compliant municipalities before and after grant receipt with equation 3.10,

$$\begin{aligned} R_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * compliant_i \\ + \sum_{y=0}^{11} \gamma_y 1\{t - t_i^* = y\} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it} \end{aligned} \tag{3.10}$$

then summarize this event study with equation 3.8.[47]

Table 3.8: Non-Compliant Cities are Valid Counterfactuals for Compliant Cities

	(1) Grant Amount	(2) Sewerage Capital	(3) Water Operations Cost
Grant X Compliant	-0.501	-4.786*	-0.513
	(2.967)	(2.450)	(0.907)
Observations	4752	4752	4752

Standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of θ from $Y_{it} = \alpha_0 + \theta grant_{it} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it}$. The dependent variable is grant amount in column 1, sewerage capital spending in column 2, and water operations costs in column 3.

Since we only rely on comparisons between compliant and non-compliant municipalities, the identifying assumption of this design is that, if not for the differences in compliance, average water revenues would have trended similarly in compliant and non-compliant municipalities after grant receipt. Table 3.8 supports comparing spending outcomes between compliant and non-compliant municipalities. Column 1 estimates equation 3.8 with grant amount as the dependent variable. The result is small and insignificant, which shows that compliant and non-compliant municipalities receive similarly sized grants. Column 2 re-estimates equation 3.8 with sewerage capital spending as the dependent variable. Even though grants to compliant and non-compliant municipalities are similar in size, per capita sewerage capital spending is $4.79 lower in compliant municipalities than in non-compliant municipalities after grant receipt. In column 3, we check if changes in revenue are driven by falling costs in compliant municipalities by estimating equation 3.8 with water operations costs as the dependent variable. The coefficient is small and insignificant, indicating that grants are not associated with changes in the cost of operating water utilities between compliant and non-compliant municipalities

[47] In our crowd-out event studies, we report coefficients for 16 years before and 11 years after grant receipt, which allows us to report only balanced coefficients. These specifications also includes bins for 17 or more years before grant receipt and 12 or more years after grant receipt, but our results are not sensitive to this choice of binning. We examine how effects evolve in later years in Appendix Section 3.7.2.3.

3.5.2 Redistribution Results

Figure 3.7 plots the π_y and γ_y from estimating equation 3.10 with per capita water revenue as the dependent variable. The null estimates in the pre-treatment period provide evidence of parallel trends in water revenues in compliant and non-compliant municipalities prior to grant receipt. The estimates begin to decrease after grant receipt and continue to fall for 11 years after treatment. The gradual decrease in water revenue is consistent with municipalities receiving multiple grants that they could spend over several years.

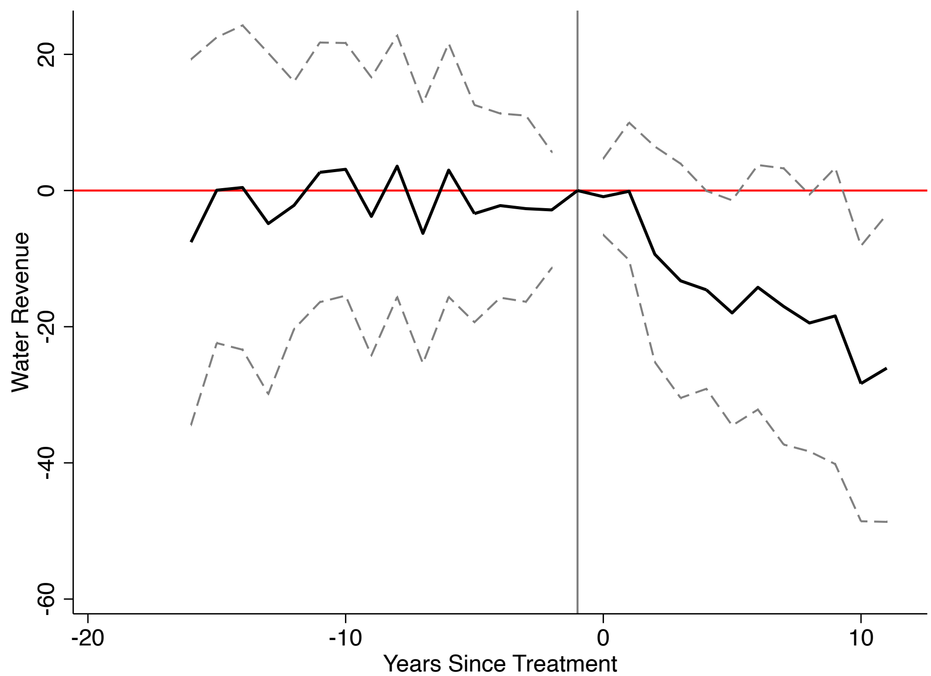


Figure 3.7: Water Revenue Decreased After Grant Receipt in Compliant Municipalities

Notes: This figure plots the π_y and γ_y from estimating $R_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * compliant_i + \sum_{y=0}^{11} \gamma_y 1\{t - t_i^* = y\} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it}$ with per water revenue as the dependent variable. $compliant_i$ equals one for observations of compliant municipalities. We include municipality and timing-group-by-year fixed effects, α_i and α_{gt}, where g indexes the year in which a municipality received its first CWA grant.

Column 1 of Table 3.9 summarizes this effect by estimating equation 3.8 with water revenue as the dependent variable. This estimate shows that per capita water revenue decreases by \$2.32 per person in compliant municipalities after grant receipt relative to non-compliant municipalities that receive grants in the same year. Compared to our sample's median per capita water revenue of \$23, this estimate represents a 10 percent decline in water revenues raised by compliant municipalities in response to grant receipt.[48]

Columns 2-4 of Table 3.9 present alternative specifications. The shape of the event study in Figure 3.7

[48] Figure 3.11 checks if this effect is consistent across timing groups by presenting the θ_g^{2X2} from estimating equation 3.9 on each timing group separately. All of the estimates in this figure are negative and relatively close to the result in column 1 of Table 3.9, so no particular within-timing group comparison is driving this result. This figure does not include coefficients for the 1980, 1985 or 1988 treatment cohorts because either no compliant or no non-compliant units were treated in those years. We omit the coefficient for the 1978 cohort because only one compliant and one non-compliant municipality became treated in that year.

suggests that this figure might be better summarized with a trend-break specification, so column 2 estimates a version of equation 3.9 where we interact the treatment dummy with year relative to treatment . This estimate suggests that water revenues decrease by an additional $0.36 each year after treatment. The map in Figure 3.2 shows that the estimates in columns 1 and 2 of Table 3.9 rely on comparisons across large geographic areas, so in column 3, we re-estimate the specification in column 2 with region-by-year fixed effects.[49] This estimate, identified off of within-region variation in compliance, is not meaningfully different than the estimate in column 2, and our result is still strongly significant despite losing a substantial amount of variation. Finally, column 4 shows that our results are robust to including year fixed effects in place of cohort-by-year fixed effects.[50]

Table 3.9: Compliant Cities Spent Crowded Out Funds on Lowering Water Bills

	(1) Water Revenue	(2) Water Revenue	(3) Water Revenue	(4) Water Revenue
Grant X Compliant	-2.323**			
	(1.157)			
Grant X Compliant X e		-0.364***	-0.396***	-0.441***
		(0.138)	(0.150)	(0.132)
Timing Group X Year FE	X	X	X	
Region X Year FE			X	
Observations	4752	4752	4752	4840

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table estimates the reduced form effect of grants to compliant facilities on water revenues. Column 1 presents estimates of θ from $R_{it} = \alpha_0 + \theta grant_{it} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it}$. Column 2 presents estimates of θ from $R_{it} = \alpha_0 + \theta grant_{it} * compliant_i * e_{it} + \alpha_i + \alpha_{gt} + \varepsilon_{it}$ where e_{it} indicates year relative to treatment. Column 3 re-estimates the result in column 2 with region-by-year fixed effects. Column 4 re-estimates the result in column 2 with year fixed effects instead of timing-group-by-year fixed effects.

These results show that water revenues decreased in compliant municipalities relative to non-compliant municipalities after grant receipt. This suggests that compliant municipalities redistributed grant money to residents by lowering water utility prices, consistent with public utility providers offsetting water and sewerage utility costs with CWA grant money in order to lower prices to consumers.

3.6 Discussion & Conclusion

There was substantial heterogeneity in the effect of CWA grants on sewerage capital spending. Municipalities spent grant money on capital upgrades up to the point where they were in compliance with the CWA's new treatment technology requirements, but after municipalities met these requirements, or if none were in place,

[49] We divide the United States into four regions; the Northeast, consisting of CT, MA, NH, NJ, NY, PA and RI, the Midwest, consisting of IA, IL, IN, KS, MI, MN, MO, NE, OH and WI, the South, consisting of AR, AL, DE, FL, GA, LA, MD, MS, NC, OK, SC, TN, TX, VA and WV and the West consisting of AZ, CA, CO, MT, NM, OR, UT and WA. States not listed do not appear in our data.

[50] Figures 3.12 and 3.13 present the event studies associated with columns 3 and 4 of Table 3.9.

grants crowded out money that municipalities were already spending on sewerage capital. Per capita water revenues decreased by ten percent from the median after grant receipt in municipalities with low pass-through relative to municipalities with higher pass-through, suggesting that municipalities redistributed crowded out funds to residents through reductions in water bills. In total, receiving governments in our sample spent 44.8% of all CWA grant dollars on sewerage capital.

Given previously quantified benefits of the CWA, accounting for low pass-through can lead to favorable benefit to cost ratios of CWA-funded capital upgrades. Dividing previously estimated benefit/cost ratios of CWA grants by our pass-through estimates suggests that the CWA grant dollars that municipalities spent on sewerage capital have a benefit/cost ratio of 1.01. This result comes from an analysis of $25.5 billion of the $153 billion in grants funded by the CWA. While our sample does not include grants distributed to relatively small municipalities, pass-through is not strongly correlated with population in our sample, which suggests that CWA grants distributed to smaller municipalities have a similar pass-through rate.

Heterogeneity in pass-through informs which types of variation researchers can use to recover the effect of the average grant dollar on spending. Many evaluations of intergovernmental grant programs estimate pass-through using difference-in-differences designs with continuous treatment, but when the effect of a grant of a given size is not constant across receiving governments, estimates identified off of continuous variation in grant size that do not account for heterogeneity in pass-through across units can either over- or under-estimate the effect of the average grant dollar on targeted spending. In the case of the Clean Water Act, estimates of grant pass-through identified off of variation in grant size over-estimate the total pass-through rate.

The relationship between grant amount and targeted spending is likely heterogeneous for many grant programs. As an example, the American Recovery and Reinvestment Act of 2009 distributed almost $50 billion to state departments of education to offset budget shortfalls caused by the Great Recession through the State Fiscal Stabilization Fund. The costs that state governments faced depended on the severity of the local economic downturn, but the federal government determined grant amount based on state population (Superfine, 2011). Budget shortfalls varied widely across states, so if pass-through depended on the costs that receiving governments faced to enact policy, it is unlikely that all states would have responded to transfers of a given size in the same way. More generally, grant size is often positively correlated with receiving governments' preferences for public goods (Knight, 2002), so the effect of a dollar of grant funding on targeted spending is likely increasing in grant size in many settings.

We cannot know for sure if heterogeneity in the relationship between grant amount and targeted spending is the cause of the inconsistent estimates of grant pass-through found in other settings, but future research that accounts for heterogeneity in pass-through across receiving governments (as well as heterogeneity across time relative to treatment in cases where the timing of grant receipt varies) may provide further evidence that this is the case. That being said, researchers are usually limited in the types of variation they can leverage to estimate pass-through; donor governments often distribute grants in such a way that there is no never-treated group and no variation in treatment timing, making it difficult to test for homogeneity in pass-through across receiving governments.

We conclude by noting several implications of this study with respect to the design of intergovernmental grant programs. First, crowd-out may be high in our setting because the CWA allotted grants toward a function of local government that had a direct revenue stream, enabling grant dollars to "drip" from where they hit down to local residents through the low-friction adjustment of water rates. This suggests that crowd-out may be more severe in settings where local governments have a low-friction channel through which to redistribute grants. Key differences between grants that target environmental spending and other intergovernmental grant programs may limit the generalizability of this result; since downstream communities are the primary beneficiaries of environmental investment (e.g. communities downstream from a wastewater treatment plant or communities downwind from air polluters), a city's environmental spending is not salient to constituents, whereas fees (in this case, water bills) are. In many other settings, taxpayers observe their state or local government receiving federal grants for a specific purpose and would notice if targeted spending did not increase, so in most cases both spending and fees are salient to constituents.

Second, pairing grants with new regulation creates a binding constraint, which (weakly) increases the likelihood of full pass-through and the desired policy outcome. Many grant programs use "maintenance of effort" requirements to achieve a similar end, but these requirements are notoriously difficult to enforce. In essence, new regulation can act like a binding maintenance of effort requirement that is relatively easy to enforce, since a donor government can check if a receiving government satisfies a mandate rather than auditing that receiving government's spending. For this method of enacting policy to be effective, donor governments must be able to identify which receiving governments need fiscal assistance and accurately estimate the costs of grant-funded programs for each receiving government.

Similarly, our results suggest that supplementing new regulation with grants can allow local governments to comply with new requirements without reducing spending in other categories (Baicker, 2001; Baicker and Gordon, 2006) or raising additional revenue (Jerch, 2018).

3.7 Appendix

3.7.1 Additional Pass-Through Results

3.7.1.1 Selection on Treated Potential Outcomes

In Section 3.4.4.2, we document heterogeneity in grant pass-through by grant size and compliance with the CWA's treatment technology mandate, but these are not necessarily the only sources of heterogeneity in pass-through. Since states distribute CWA grants to municipalities according to priority lists, we might expect treatment timing to be positively correlated with treated potential outcomes. This is not a threat to identification (we do not place any restrictions on the evolution of outcomes after treatment and our identifying assumption allows for selection into treatment on the basis of unobserved time-invariant characteristics), or the interpretation of the heterogeneity we document in Section 3.4.4.2 (as shown in Figures 3.18 and 3.19, the distribution of treatment timing is similar across compliant and non-compliant municipalities), but it does have implications for the best way to summarize treatment effects. Callaway and Sant'Anna (2020b) suggests that when treatment effects vary across timing groups, we should first aggregate the $ATT(g,t)$ for each timing group g, then combine these timing group-specific effects into a summary measure by taking a weighted average of the effect in each timing group. Specifically, we first calculate

$$\tilde{\theta}(g) = \frac{1}{1985 - g + 1} \sum_{t=1972}^{1987} 1\{g \leq t\} ATT(g,t)$$

for each group, then summarize the effect with

$$\theta_S = \sum_{g=1972}^{1985} \tilde{\theta}(g) P(G = g)$$

θ_S is similar to the simple aggregation based on group-size given by equation 3.2. The difference is that the weights in θ_S do not depend on how long a unit has been treated for, while the weights in equation 3.2 depend on how many periods we observe a timing group for after treatment. This weights up groups that are treated early, so in the presence of selection on gains, equation 3.2 weights up units that experience the largest effects. In contrast, the weights in θ_S only depend on group size.

Table 3.17 presents averages of the $\tilde{\theta}(g)$ for all municipalities in the first three treatment groups, the second three treatment groups and the last five treatment groups in columns 1-3. Column 4 presents θ_S for the entire sample. θ_S is similar to the dynamic and simple aggregations in Panels A and D of Table 3.2, which suggests that treatment timing is not correlated with treated potential outcomes.

Column 5 calculates θ_S with grant amount as the dependant variable. Dividing the result in column 4 by this first stage estimate implies a pass-through rate of 0.46, which is consistent with the results in Table 3.2.

3.7.1.2 Alternative Definitions of Full Pass-Through

Cost-sharing yields two different possible interpretations of pass-through. We define full pass-through as \$1 in grant funding leading to a \$1 increase in sewerage capital spending. Alternatively, since municipalities are required to pay 25% of project costs, one could define full pass-through as \$1 in grant funding leading to a \$1.33 increase in sewerage capital spending (the actual \$1 in grant aid plus the \$0.33 in local obligations). Our estimates in the main text are based on the first definition, appealing to the notion of federal funding itself being spent as intended. We present pass-through estimates using the second definition in Tables 3.18 and 3.19. These tables re-estimate the results in Tables 3.2 and 3.4 after multiplying grant amount by 1.33 in years before 1981, and by 1.82 in 1981 and beyond. Using this definition of full pass-through, we still reject full pass-through for the full sample, non-compliant municipalities that received large grants, and compliant municipalities. We cannot reject full pass-trough or complete crowd-out for non-compliant municipalities that received small grants, but the point estimate for these municipalities suggests that, while CWA grants led to an increase in targeted spending equal to grant amount, they did not cause an increase in local contributions to sewerage capital spending.

3.7.1.3 Average Pass-Through Rate

In the main text, we attempt to recover the total pass-through rate of CWA grants, but that is not the only potential parameter of interest. Figure 3.20 presents the distribution of per capita grant amount for grants in our sample and shows that grant amount is distributed approximately lognormally; given this distribution, it is possible that a few larger grants are meaningfully driving our estimates of the total pass-through rate. Since we want to place the largest weight on the most expensive grants when analyzing the effectiveness of a grant program, this is a desirable feature. That being said, we may instead be interested in the "average pass-through rate", the percent of grant dollars spent as the donor government intended in the average receiving municipality.

We can estimate the average pass-through rate by estimating pass-through in each unit separately, then taking an average across municipalities. To do this, we first estimate equations 3.4 and 3.11 on observations in stack each S.

$$g_{it} = \alpha_0 + \beta_{s_{fs}} post_t * treat_i + \delta post_t + \omega treat_i + \varepsilon_{it} \tag{3.11}$$

We then divide β_s^{2X2} by $\beta_{s_{fs}}$ to obtain an estimate of pass-through for municipality s, then average the resulting within-stack pass-through rates across municipalities.

We present estimates of the average pass-through rate in Table 3.20. While we cannot reject full pass-through or full crowd-out, the point estimate is similar to our estimates of the total pass-through rate.

3.7.1.4 Excluding Municipalities that did not Receive Grants

Since grant timing is determined by a municipality's placement on a state priority list and priority lists are explicitly based on factors that are unlikely to be correlated with trends in sewerage capital spending, we identify our main results off of variation in the *timing* of grant receipt. It is less clear if grant *receipt* is uncorrelated with trends in local spending, but if we are willing to assume that whether or not a municipality received a grant is uncorrelated with trends in local spending, we can use municipalities in our finance data that never receive a CWA grant as an additional control group.

Using never-treated units as a control group only requires minor changes to our estimators. Our Callaway and Sant'Anna (2020a) estimator works in exactly the same way, except we now estimate group-time average treatment effects with

$$\widehat{ATT(g,t)} = E_n[\widehat{C}_t - \widehat{C}_{g-1}|N_i = 1] - E_n[\widehat{C}_t - \widehat{C}_{g-1}|N_i = 0]$$

where N_i is a dummy variable that equals one for never treated units. We can then aggregate these group-time treatment effects together the same way we did in Section 3.4. We present these results in Figures 3.14 and 3.15 and Table 3.21. These estimates are similar to our main results.

For our stacked estimator, we construct stacks so that, instead of using units treated at least seven years after the treated unit as controls, we use never-treated units as controls. We present these results in Table 3.22.

3.7.1.5 Observable Characteristics do not Predict Grant Timing

The identifying variation of our grant pass-through research design comes from *when* each municipality receives its first CWA grant. To identify causal effects with this variation in a difference-in-difference framework, we must assume that the timing of grant receipt is uncorrelated with trends in sewerage capital spending. We perform more explicit tests of our parallel trends assumption in the main text, but in this section, we provide suggestive evidence that early and late treated units would have followed parallel trends in absence of treatment by showing that early and late treated municipalities are observably similar before treatment. We test if observable characteristics of municipalities predict the timing of grant receipt with equation 3.12.

$$GrantYear_i = \alpha_0 + \alpha_1 compliant_i + \alpha_2 X_{ip} + \varepsilon_{ip} \tag{3.12}$$

We regress the year municipality i receives its first CWA grant on observable municipality characteristics

X from a baseline year p, conditional on the municipality not yet receiving a grant in year p. Table 3.23 presents results from estimating equation 3.12 for baseline years 1971, 1974 and 1977.[51] X_{ip} includes city population and per-capita measures of both holistic (e.g. total revenue, total intergovernmental revenue, total capital outlays) and specific (e.g. sewerage capital outlays, water utility revenue) categories of government revenue and expenditure.

None of these observable characteristics consistently predict grant timing. All of the point estimates are small and precisely estimated and the only covariates that have a significant effect on year of grant receipt are water capital spending and log population using 1971 as a baseline year. The coefficients on these variables do not represent economically meaningful changes; a ten percent difference in 1971 population is associated with a municipality receiving its first grant 0.08 years earlier and a one standard deviation increase in water capital outlay is associated with a less than one year change in grant timing. Both of these characteristics are no longer significant and change signs when we use other baseline years. While the covariates are jointly significant when we use baseline years of 1971 and 1974, a standard deviation change in any spending variable in either year represents a less than one year change in grant timing. This shows that grant timing is not strongly correlated with any municipal spending characteristics, which motivates a design that leverages variation in treatment timing to estimate pass-through.

3.7.1.6 Alternative Pass-Through Specifications

In Section 3.4, we summarize pass-through for the full sample using Callaway and Sant'Anna (2020a) and explore heterogeneity in pass-through using a stacked difference-in-differences estimator. We discuss our choice of estimators in greater detail in Appendix 3.7.3, and in this section, we present additional results to show that our results not sensitive to this choice.

Figure 3.16 presents the π_y and γ_y from estimating the following stacked event study on a pooled sample of all stacks.

$$C_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * treat_{is} + \sum_{y=0}^{6} \gamma_y 1\{t - t_i^* = y\} * treat_{is} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its} \tag{3.13}$$

Time relative to treatment is defined by $1\{t - t_i^* = y\}$, which is a dummy variable that equals 1 for observations of municipality i y years before or after t_i^*, the year municipality i receives its first CWA grant and $treat_{is}$ is a dummy variable that equals one for the treated unit in each stack. Similar to the results in Figure 3.4, we see null effects before treatment. The estimates increase after the arrival of a municipality's first CWA grant and remain high for up to 6 years after grant receipt.

[51] Roughly one third of grants to municipalities in our sample were distributed from 1972 to 1973, 1974 to 1976 and post-1976.

Table 3.24 summarizes the results in Figure 3.16. Column 1 estimates the first stage relationship between grant receipt and per-capita grant amount by estimating equation 3.5 on a pooled sample of all stacks. Column 2 estimates the associated reduced form effect of grant receipt on sewerage capital spending by estimating equation 3.6 on all stacks. Column 3 estimates equation 3.6 using D_{cy} as an instrument for grant amount, which returns an estimate equivalent to dividing the reduced form by the first stage. This implies a pass-through rate of 0.342. Using this result, we can reject full pass-through in the full sample (i.e. a test of the hypothesis that the IV estimate in column 3 equals one).

Table 3.25 re-estimates the results in Panel B of Table 3.2 on sub-samples of municipalities. These results are averages of the $\widehat{\theta}_D(e)$ from equation 3.1 for e s.t. $0 \leq e \leq 6$. Overall, these results are similar to those from our stacked difference-in-difference; we can reject full pass-through in non-compliant municipalities that received large grants, but cannot reject full pass-through for non-compliant municipalities receiving small grants. Since the effect of grant receipt on sewerage capital spending is marginally significant for non-compliant municipalities receiving small grants, we can also marginally reject complete crowd-out in these municipalities.

We obtain similar results in Table 3.26, where we re-estimate the results in Table 3.25 using the aggregation given by equation 3.2 (equivalent to Panel D of Table 3.2).

We cannot say anything about compliant municipalities using this estimator. Since there are relatively few compliant municipalities in our sample, the simultaneous critical value is arguably 'too large' to be reliable when we estimate the effect of grant receipt on sewerage capital spending in these municipalities, so we default to our stacked results.

3.7.1.7 Measuring the Benefits of CWA Grants

In this section, we discuss several potential issues with using previously estimated benefits of CWA grants to find the benefit-to-cost ratio of CWA grants that municipalities spent on sewerage capital.

First incorporating other currently unmeasured benefits could further increase the benefit-to-cost ratio CWA grants, and a full welfare analysis would include these unmeasured benefits.

Second, a more formal analysis would take the confidence intervals and potential bias of estimates of the benefits of CWA grants into account. In particular, both the hedonic estimates in Keiser and Shapiro (2019) and the infant health estimates in Flynn and Marcus (2021) come from TWFE estimators with continuous treatment, but we show in Section 3.4.5 that this type of estimator can produce misleading results in the presence of heterogeneity in treatment effects across compliance or heterogeneity in treatment effects across time. The main estimates in Flynn and Marcus (2021) account for heterogeneity across compliance. Flynn and Marcus (2021) also obtains similar results when using estimators that account for dynamic treatment effects. While it does not explicitly account for heterogeneity across time, the event study graphs in Keiser and Shapiro (2019) do not show evidence of dynamic treatment effects, and instead show that CWA grant receipt led to a mean shift in housing prices. Keiser and Shapiro (2019) does not account for heterogeneity across compliance either, but since there is evidence that similar demographic groups sorted into communities downstream from compliant and non-compliant municipalities (see Table 2 in Flynn and Marcus (2021)), it is not clear that the effect of CWA grants on housing prices is heterogeneous across compliance. That being said, if this heterogeneity is present, it could bias estimates of the effect of CWA grants on housing prices. It is not clear a priori which direction this bias would move in.

Finally, the benefits of CWA grants may be driven in part by spending on categories besides sewerage capital. In Section 3.5, we provide evidence that compliant municipalities redistributed grant dollars to residents through reduced water bills. A hedonic model suggests that these savings should be capitalized into housing prices. Since Keiser and Shapiro (2019) estimates the effect of CWA grant receipt on housing prices by comparing all housing units within 25 miles of treated plants against all housing units within 25 miles of untreated plants, the estimated benefit-to-cost ratio of 0.26 in Keiser and Shapiro (2019) may be driven in part by the value that residents place on reduced water bills. Consequently, our estimated benefits likely include the value that residents place on lowered water rates in addition to the benefits of improved infant health and the value that residents place on increased water quality.

Reductions in water rates are driven by grant dollars that municipalities redistributed to residents. For this reason, we do not want to include these benefits when we are trying to find the benefit-to-cost ratio of CWA grant dollars spent on sewerage capital, but we cannot disentangle the portion of the effect of CWA grants on housing prices driven by lower utility prices and the portion driven by improved water quality. We can instead find the benefit-to-cost ratio of CWA grant dollars spent on sewerage capital *and* CWA grant dollars that municipalities redistributed to residents through lowered water bills. Table 3.32 suggests that 48.5% of grant dollars for compliant facilities went to reducing water bills (though this estimate is imprecise). Twelve percent of CWA grant dollars went to compliant municipalities, so this represents 5.82% of all CWA grant dollars. Since the benefit-to-cost ratio of 0.45 from Keiser and Shapiro (2019) and Flynn and Marcus (2021) comes from both the 44.6% of CWA grant dollars that municipalities spent on sewerage capital and the 5.82% of all CWA grant dollars went to reducing water bills, we can obtain the benefit-to-cost ratio of CWA grant dollars that were spent on sewerage capital or redistributed to residents through reduced water bills by dividing previously estimated benefit-to-cost ratios of CWA grants, 0.45, by the portion of CWA grant dollars spent on sewerage capital and reducing water bills combined, 0.504. This suggests a benefit-to-cost ratio of 0.89, however this is not exactly the number that we are interested in.

We want the benefit-to-cost ratio of CWA grant dollars spent on sewerage capital. We can use the results in Table 3.32 to estimate a range for this number. To generate a lower bound, we can assume that reductions in water rates were completely capitalized into housing prices (i.e. a one dollar reduction in water bills leads to a one dollar increase in housing prices). If this were the case, then, of the \$0.26 of the average dollar of CWA grant aid capitalized into housing prices, 0.058 would come from reduced water bills and the other 0.202 would come from improvements in water quality. Combining this number with the infant health benefits from Flynn and Marcus (2021) suggests that the CWA grant dollars spent on everything except reducing water bills would have a benefit-to-cost ratio of 0.392. We can divide this number by the pass-through rate to calculate the benefit-to-cost ratio of CWA grant dollars spent on sewerage capital. This suggests a benefit-to-cost ratio of about 0.879. If instead reductions in water bills were not capitalized into housing prices at all, all of the benefits of the CWA in Keiser and Shapiro (2019) and Flynn and Marcus (2021) would come from sewerage capital spending, and the benefit-to-cost ratio of CWA grant dollars spent on sewerage capital spending would be $\frac{0.45}{0.446}$, or about 1.01. Where in this range the actual benefit-to-cost ratio of CWA grant dollars spent on sewerage capital spending falls depends on the capitalization rate of reductions in water bills, but the existing literature has not reached a consensus on the effect of changes in fees on housing prices (Yinger et al., 2016). Recent results suggest that there should be near full capitalization if the housing supply is inelastic (Lutz, 2015), but determining the elasticity of housing around municipal wastewater treatment facilities is outside the scope of this paper.

Similarly, the portion of CWA grant dollars that municipalities spent on categories besides sewerage capital spending and lowering water bills may have caused increases in housing prices or improvements in infant health, which would similarly decrease the benefit-to-cost ratio of CWA grants spent on sewerage capital.

3.7.1.8 Baseline Sewerage Spending by Compliance and Grant Size

If municipalities that received small grants had relatively low pre-treatment sewerage capital spending, it could be the case that a relatively small portion of the grants distributed to these municipalities are inframarginal. If less of these grants dollars are inframarginal, then the higher pas-through that we document in municipalities that receive small grants could be mechanical; municipalities that received small grants would not have had much spending to to crowd-out, while municipalities that received larger grants spend more in sewerage capital before grants, and would have had more baseline spending to crowd out. Table 3.27 shows that this is not the case, and that pre-treatment spending is highest in non-compliant municipalities that receive small grants.

3.7.2 Additional Water Revenue Results

3.7.2.1 Estimating Water Revenue Results Using Timing Variation

In Section 3.5, we leverage variation in pre-CWA compliance with the CWA's treatment technology mandate to examine how compliant municipalities adjust spending in response to grant receipt, but we obtain similar results when we estimate these effects using variation in treatment timing.

In Figures 3.21 and 3.22, we estimate separate event studies for compliant and non-compliant municipalities. These event studies compare water revenue after treatment to water revenue in not-yet-treated municipalities using Callaway and Sant'Anna (2020a). Sub-Figure 3.21 shows that grant receipt has no significant effect on water revenues in non-compliant municipalities. Sub-Figure 3.22 shows that water revenues are stable before grant receipt and decrease after grant receipt in compliant municipalities. Table 3.28 summarizes these figures.

Note that the estimates in column 1 of Table 3.28, which are identified off of comparisons between early and late treated compliant municipalities, are larger than the estimates in Table 3.9, which are identified off of comparisons between compliant and non-compliant municipalities. Similarly, the event study in Figure 3.22 has a somewhat different shape than the event study in Figure 3.7. This does not reflect a problem with either design. Rather, it demonstrates that these two sources of variation identify different treatment effects.

This difference is caused in part by the way in which OLS combines the effects in different treatment cohorts into summary measures. OLS weights based on the size of timing groups and within-group variance of the treatment dummy, which artificially weights up units treated near the middle of the panel (Goodman-Bacon, 2021a). To avoid weighting based on within-group variance, we can take a weighted average of the effect in each timing group where the weights are based only on the size of each timing group (i.e. a weighted average of the effects in Figure 3.11). This average is -3.76, which is closer to the estimate in column 1 of Table 3.28.

3.7.2.2 Why was there No Effect in Non-Compliant Municipalities?

Table 3.28 shows that grant receipt had no average effect on water revenue in non-compliant municipalities. This result provides support for the identifying assumption of the design we use in Section 3.5, but it is somewhat counter-intuitive; since we do not estimate full pass-through in all non-compliant municipalities in our sample, we might expect grants to cause a decrease in water revenues in non-compliant municipalities.

Figure 3.1 suggests one explanation for this result. Note that while, on average, small grants are similar to costs and large grants are greater than costs, there are non-compliant municipalities that receive grants that are too small or too large along the entire distribution of grant size. If this leads some municipalities redistribute grant funds by decreasing their water bills while other municipalities increase water bills to cover costs greater than grant amount, these effects will cancel out.

Figures 3.23 and 3.24 provides suggestive evidence that this is the case by re-estimating the results in Figure 3.21 on sub-samples of municipalities where grants are larger or smaller than estimated costs. Figure 3.23 suggests that municipalities that receive grants that are smaller than costs increase their water bills in order to raise the revenue needed to come into compliance with the CWAs treatment technology mandate. Figure 3.24 does not provide clear evidence of an increase or decrease in water revenues in non-compliant municipalities that receive grants larger than estimated costs.

Table 3.29 summarizes the results in Figures 3.23 and 3.24. The estimate for municipalities that receive grants smaller than costs is positive, while the estimate for municipalities that receive grants larger than costs is negative, as expected, though neither of these effects are statistically significant. This suggests that grants did not cause a decrease in water revenues in non-compliant municipalities because these countervailing effects cancel out.

The results in Table 3.28 also motivate a discussion of how our results relate to Jerch (2018). Jerch (2018) uses an instrumental variables approach to estimate the effect of the CWA's capital mandate on local spending and finds that the capital mandate imposed a large financial burden on non-compliant municipalities, and that non-compliant municipalities then increased water bills in response to the capital mandate. Two important differences between this paper and Jerch (2018) can explain why we obtain different results when we estimate the effect of grant receipt on water revenue for non-compliant municipalities: first, the research design in Jerch (2018) intentionally does not capture the effect of CWA grants, and second, the sample of municipalities that Jerch (2018) uses includes municipalities that did not receive any CWA grants. In contrast, we only look at the effect of the CWA's capital mandate on water revenues in municipalities that received CWA grants, and find that, on average, non-compliant municipalities that received CWA grants did not change their water bills. This suggests that the increase in water revenues in non-compliant municipalities documented in Jerch (2018) are driven by municipalities that did not receive CWA grants. This is consistent with other studies of unfunded federal mandates (Baicker, 2001; Baicker and Gordon, 2006).

3.7.2.3 How Long Do Effects Persist?

Figures 3.7 and3.22 both show that the effect of grants to compliant municipalities on water revenue grow over time, which motivates an examination of how long these effects persist. Our spending data includes observations from 1956 to 1999 and the latest treated municipality receives its first CWA grant in 1988. This allows us to estimate event study coefficients for 12 balanced post-treatment periods in our main specifications, but if we drop the compliant municipalities treated after 1981, we can estimate event study coefficients for 19 balanced post-treatment periods.

Figure 3.25 re-estimates the results in Figure 3.7 on this sample. The estimates for the 16 years before treatment and the first 12 years after treatment are very similar to those in Figure 3.7. The effect of grant receipt on water revenue reaches its lowest point 13 years after treatment, then begins to return to pre-treatment levels.

This pattern of effects is consistent with municipalities receiving multiple grants that were fungible across years. Figure 3.26 shows the distribution of grants that a municipality received each year relative to the year the municipality receives its first grant. This figure shows that some municipalities receive CWA grants up to 16 years after initial grant receipt, which can explain why the effect of grant receipt on water revenues persists well into the post-treatment period.

3.7.2.4 Alternative Water Revenue Data

Starting in 1967, the Census began collecting spending data from a larger sample of municipalities once every five years. Since it only contains one pre-treatment period, we do not use this data in our main crowd-out specifications, however, we obtain similar results when we re-estimate the results in Sections 3.5 and 3.7.2.1 using this dataset.

Figure 3.27 re-estimates the results in Figure 3.7 on this data. The pattern of effects is similar to those in Figures 3.7 and 3.25. Water revenue decreases in compliant municipalities relative to non-compliant municipalities for 15 years after grant receipt, then begin to return to pre-treatment levels.

Table 3.30 summarizes Figure 3.27. Column 1 estimates equation 3.8 on our alternative dataset. Column 2 re-estimates the specification in column 1 with region-by-year fixed effects, and column 3 re-estimates the specification in column 1 with a year fixed effect in place of a timing-group-by-year fixed effect. The results are similar to the estimate of equation 3.8 in column 1 of Table 3.9.

Figures 3.28 and 1.29 and Table 3.31 then re-estimate the results in Figures 3.21 and 3.22 and Table 3.28 respectively. The results are again similar to those we obtain with our main dataset.

3.7.2.5 How Much Crowd-Out can Redistribution Account For?

In Section 3.5, we show that, relative to non-compliant municipalities that received grants in the same year, compliant municipalities reduced water bills by \$2.32 per capita after grant receipt. In this section, we examine how much crowded-out spending reductions in water revenue account for.

Column 1 of Table 3.32 presents estimates of equation 3.8 with sewerage capital spending as the dependent variable. Panel A estimates this equation with a dummy treatment variable and Panel B re-estimates this specification interacting the treatment dummy with year relative to treatment. These results suggest that, after grant receipt, compliant municipalities increased sewerage capital spending by \$4.76 per capita (or by an additional \$0.51 per year) less than non-compliant municipalities that were treated in the same year. Since, as shown in column 1 of Table 3.8, compliant and non-compliant municipalities that were treated in the same year received grants of similar size, this difference in sewerage capital spending is likely due to differences in pass-through. For this reason, dividing our reduced form estimates of the effect of grant receipt on water revenues in compliant municipalities from Table 3.9 (also presented in column 3 of Table 3.32) will show how much crowded-out spending reductions in water revenue can account for. We present these results in column 3 of Table 3.32. While they are imprecise and we cannot reject that the coefficients equal 1 (which would suggest that reductions in water revenue can account for all crowded-out spending) or 0 (which would suggest that they do not account for any crowded out spending), taken at face value, these results suggest that reductions in water revenue can account for between half and three-quarters of crowded-out spending.

Since we do not find evidence of an increase in targeted spending in compliant municipalities, any spending that this does not account for was likely redistributed to other functions of local government according to local preferences.

3.7.3 Choice of Estimators

There are many ways to estimate treatment effects in a difference-in-differences framework, including two way fixed effects estimators with binary and continuous treatment, Callaway and Sant'Anna (2020a), and stacked difference-in-difference estimators. Each of these methods identifies off of different comparisons and combines these comparisons into summary measures according to different weights. The decision of which method to use is context specific, so we rely on different estimators for different portions of our analysis. In this section, we discuss our choice of estimators in detail.

3.7.3.1 Two Way Fixed Effects with Binary Treatment

We could leverage variation in grant timing with a two way fixed effects (TWFE) estimator. To do this, we would first estimate the reduced form effect of grant receipt on sewerage capital spending with the following equation

$$C_{it} = \alpha_0 + \beta_{rf}^{TWFE} D_{it} + \alpha_i + \alpha_t + \varepsilon_{it} \tag{3.14}$$

where C_{it} is per capita sewerage capital spending, D_{it} is a dummy variable that equals one for observations of municipalities after they receive their first grant and α_i and α_t are municipality and year fixed effects. We would then divide our estimate of β_{rf}^{TWFE} by an estimate of the first stage relationship between grant receipt and per capita grant amount, g_{it}, from estimating

$$g_{it} = \alpha_0 + \beta_{fs}^{TWFE} D_{it} + \alpha_i + \alpha_t + \varepsilon_{it} \tag{3.15}$$

Since every unit in our sample is eventually treated, estimating β_{rf}^{TWFE} returns an average of comparisons between (1) newly treated municipalities relative to municipalities that have not yet been treated and (2) newly treated municipalities relative to already-treated municipalities. When treatment effects are dynamic, municipalities still *actively responding to treatment* are not a valid counterfactual to represent potential outcomes in the *absence of treatment*, so the second type of comparison can produce results that do not have a causal interpretation. Mechanically, changes in the treatment effects of already-treated units over time are subtracted from a TWFE estimate and we cannot interpret the resulting coefficient as the average treatment effect of the treated (ATT) (Goodman-Bacon, 2021a).

We have an a priori reason to expect the effect of CWA grants on sewerage capital spending to change over time. States distributed CWA grants to fund specific projects, so we would expect grant receipt to cause an increase in sewerage capital spending that only lasts until the project is completed. When spending returns to pre-treatment levels, this decrease is subtracted from a TWFE estimate, biasing the estimate upwards.

We illustrate this problem with a simple example (shown graphically in Figure 3.30). Consider a setting where there are two municipalities, A and B, and three time periods, $t = \{0,1,2\}$. Treatment turns on for municipality A in period $t = 1$ and for municipality B in period $t = 2$. After treatment turns on, the outcome variable y increases by one, then returns to pre-treatment levels in the next period. Summarizing this effect with TWFE will return an average of (1) a comparison of municipality A to municipality B in periods 0 and 1 where municipality A is the treated unit and (2) a comparison of municipality B to municipality A in periods 1 and 2 where municipality B is the treated unit. If we denote y_i^t as the outcome variable in municipality i in period t, the first comparison will be

$$(y_A^1 - y_A^0) - (y_B^1 - y_B^0) = (4-3) - (1-1) = 1$$

and the second will be

$$(y_B^2 - y_B^1) - (y_A^2 - y_A^1) = (2-1) - (3-4) = 2$$

TWFE will return a weighted average of these two comparisons, which is strictly greater than the true effect, and our result will be biased upwards. This sort of upward bias could be one reason that estimates of grant pass-through are higher than economic theory would predict in situations where grants roll out over time.

To test the validity of a TWFE estimator in this context, we check how much of our estimate of β_{rf}^{TWFE} comes from comparisons of newly treated units relative to already-treated units. We first estimate equation 3.14 on the full sample, then decompose the resulting estimate of β_{rf}^{TWFE} into every possible comparison between timing groups using Goodman-Bacon et al. (2019). Figure 3.32 presents a scatterplot of these comparisons and their associated weights. Table 3.33 summarizes Figure 3.32 and shows that more than half of our TWFE estimate comes from comparisons of newly treated units to already-treated units, raising concerns over its interpretation and motivating a different estimator.

3.7.3.2 Dose-Response Two Way Fixed Effects

Researchers often estimate pass-through by leveraging continuous variation in grant size. **?** documents several potential problems that arise when using this type of variation in a difference-in-differences framework. In this section, we discuss how the results from **?** apply to estimating grant pass-through.

Consider the following equation, which regresses targeted spending on a continuous measure of grant size and unit and time fixed effects:

$$C_{it} = \alpha_0 + \beta_{dr}^{TWFE} g_{it} + \alpha_i + \alpha_t + \varepsilon_{it} \tag{3.16}$$

When there is no never treated group, then, similar to a TWFE estimate with binary treatment, estimating β_{dr}^{TWFE} returns a weighted average of (1) comparisons of newly treated municipalities relative to municipalities that have not yet been treated, (2) comparisons of newly treated municipalities relative to already-treated municipalities, *and* (3) comparisons of municipalities that received larger grants relative to municipalities that received smaller grants. As in Section 3.7.3.1, the first type of comparison produces meaningful estimates under parallel trends. In the presence of dynamic treatment effects, the second type of comparison can produce results that do not have a causal interpretation.

To identify meaningful effects with the third type of comparison, we must assume that the effect of a grant of a given size is homogeneous across receiving municipalities. To see why, consider two groups of municipalities treated in the same period. Municipalities in one group receive grants of size d and municipalities in the the other group receive grants of size d'. Express the effect of d on the group that received grants of size d as $ATT(d|d)$. Under parallel trends, comparing the effects of grants in these groups of municipalities yields the following result:

$$ATT(d|d) - ATT(d'|d') = \underbrace{ATT(d|d) - ATT(d'|d)}_{\text{effect of going from d' to d}} + \underbrace{ATT(d'|d) - ATT(d'|d')}_{\text{selection bias}}$$

We can interpret the first term as the effect of increasing grant size from d' to d on units that actually received d, which is a causal parameter. The problem is that this type of comparison also includes a selection bias term, which will be averaged into our estimate of β_{dr}^{TWFE}.

Under parallel trends, we can write the selection term as

$$E[Y_t(d') - Y_{t-1}(0)|D = d] - E[Y_t(d') - Y_{t-1}(0)|D = d']$$

which is equal to the effect of a grant of size d' on units that actually received d minus the effect of a grant of size d' on units that actually received d'. If we assume that the effect of d' is homogeneous, then these terms cancel.

Since grants of a given size often have different effects in compliant and non-compliant municipalities, this assumption is unlikely to hold in the context of CWA grants. The following example illustrates why heterogeneity in grant pass-through by compliance presents a problem for estimation. Consider what happens when we compare a non-compliant municipality that received a grant of \$210 and increased spending by \$75 to a compliant municipality that received a grant of \$200 and increased spending by \$25. The total pass-through rate of the grants to these two municipalities is 0.244, but comparing these two municipalities implies that each dollar of grant money led to a $\frac{75-25}{210-200} = 5$ dollar increase in targeted expenditure. Comparisons of compliant municipalities to non-compliant municipalities will face the opposite problem; this type of comparison suggests that grants led to a *decrease* in spending, which will bias β_{dr}^{TWFE} downward.

We can address problems caused by heterogeneity across compliant and non-compliant municipalities by estimating β_{dr}^{TWFE} on each of these groups separately, (we present these results in Table 3.34) however, these results may not return the total pass-through rate if treatment effects are dynamic. We illustrate how comparisons across timing groups complicate estimation with a modified version of the example in Figure 3.30 (shown graphically in Figure 3.31). There are two municipalities, A and B, and three time periods, t = $\{0,1,2\}$. Treatment turns on for municipality A in period $t = 1$ and for municipality B in period $t = 2$. Municipality A receives a dose of size 2 and municipality B receives a dose of size 1. After treatment turns on, the outcome variable y increases by dose size before returning to pre-treatment levels in the next period. A TWFE estimate will include (1) a comparison of municipality A to municipality B in periods 0 and 1 where municipality A is the treated unit, divided by the dose size in municipality A, (2) a comparison of municipality B to municipality A in periods 1 and 2 where municipality B is the treated unit, divided by the dose size in municipality B and (3) a comparison of municipality A to municipality B in periods 0 and 3 where municipality A is the treated unit, divided by the the difference in doses between municipality A and municipality B. If we denote the dose size in municipality i as d_i, the first comparison will be

$$\frac{(y_A^1 - y_A^0) - (y_B^1 - y_B^0)}{d_A} = \frac{(5-3)-(1-1)}{2} = 1$$

the second will be

$$\frac{(y_B^2 - y_B^1) - (y_A^2 - y_A^1)}{d_B} = \frac{(2-1)-(3-5)}{1} = 3$$

and the third will be

$$\frac{(y_A^2 - y_A^0) - (y_B^2 - y_B^0)}{d_A - d_B} = \frac{(3-3)-(2-1)}{2-1} = -1$$

Only the comparison of early treated units to not-yet-treated units accurately describe the relationship between dose size and the outcome variable. Note that dose-response variation returns a result that is wrong signed *even though the effect of grants is homogeneous in grant size*.[52]

To interpret results identified off of dose-response variation as the ATT when treatment timing varies, we must assume that the response to a given does in a given time period is constant across groups (**?**). Since treatment effects are heterogeneous across both time relative to treatment (as shown in Figure 3.4) and compliance (as shown in Figures 3.5 and 3.6), this assumption does not hold. For this reason, our research design does not leverage variation in grant size.

3.7.3.3 Callaway and Sant'Anna

Since it focuses on combinations of $ATT(g,t)$, Callaway and Sant'Anna (2020a) is far simpler than a TWFE estimator. In practice, $\widehat{ATT(g,t)}$, the sample analogue of $ATT(g,t)$, is equivalent to the OLS estimate from a two group, two time period difference-in-difference that compares observations from group g and observations from groups not yet treated at time t in periods t and $g-1$. That is, $\widehat{ATT(g,t)}$ is identical to an estimate of β from equation 3.17.

$$C_{it} = \alpha_0 + \alpha_1 G_g + \alpha_2 1\{T = t\} + \beta(G_g * 1\{T = t\}) + e \tag{3.17}$$

Our 12 timing groups, observed across many post-treatment periods, yield 96 different $\widehat{ATT(g,t)}$. Most of the $\widehat{ATT(g,t)}$ are identified off of relatively few observations, so instead of interpreting individual $\widehat{ATT(g,t)}$, we summarize the effect of treatment by aggregating the $\widehat{ATT(g,t)}$ together. Unlike a TWFE estimator, where

[52]See Theorem 5 in **?** for a formal discussion.

comparisons are always weighted together based on group size and within-group variance of the treatment dummy (Goodman-Bacon, 2021a), researchers can select weights appropriate to the setting. We report the dynamic aggregation given by equation 3.1 and the simple aggregation given by equation 3.2 in the main text, as well as an aggregation that allows for selective treatment timing in Appendix Section 3.7.1.1.

Callaway and Sant'Anna (2020a) constructs standard errors with a multiplier bootstrap procedure.[53] Instead of re-sampling observations, each bootstrap draw perturbs the influence function of the estimate (which measures the dependence of the estimate on each cluster in the sample). We report multiplier bootstrap standard errors for our summary measures of the first stage relationship between grant receipt and grant amount and the reduced form effect of grant receipt on sewerage capital expenditure.

Rather than derive the influence function for our pass-through estimator, we construct standard errors for our pass-through estimates with a pair-bootstrap procedure. To do this, we take a random sample (with replacement) of municipalities and jointly estimate our first stage and reduced form estimates. We then divide our reduced form estimate by the associated first stage estimate to obtain β_b^{boot}, where b indexes bootstrap iterations. We repeat this process 1000 times and use the results to calculate bootstrap standard errors with the following equation

$$\sqrt{\frac{1}{999}\sum_{b=1}^{1000}(\beta_b^{boot} - \beta^{boot})^2} \tag{3.18}$$

where $\beta^{boot} = \frac{1}{1000}\sum_{b=1}^{1000}\beta_b^{boot}$.

Because pair-bootstrap draws do not preserve the relative size of each timing group, the standard errors for our pass-through estimates will not be efficient if the effect of treatment varies across treatment groups. The results in Section 3.7.1.1 suggest that this is not the case, so changes in the size of timing groups across bootstrap samples should not add variance to our pair-bootstrap estimates.[54]

3.7.3.4 Stacked Difference-in-Differences

When examining heterogeneity in the relationship between grant amount and targeted spending, we rely on estimates from a stacked difference-in-differences estimator. While most of our results are robust to using Callaway and Sant'Anna (2020a) instead, a stacked estimator allows us to define our treatment and control groups more flexibly. While Callaway and Sant'Anna (2020a) estimates the average effect of grant receipt on *all units* in a *given timing group* in *different periods*, we construct our stacks so that we estimate the average effect of grant receipt on *individual units* across *all post treatment periods*. This allows

[53] See Algorithm 1 in Callaway and Sant'Anna (2020b).

[54] Note that we do not go so far as to assume that treatment timing is truly random. When that is the case, difference-in-differences designs do not provide the most efficient estimates (Roth and Sant'Anna, 2021).

us to semi-parametrically examine the relationship between grant amount and sewerage capital spending before constructing summary measures of pass-through. Since timing groups contain both compliant and non-compliant municipalities and municipalities that received different sized grants, we cannot repeat this process with individual $ATT(g,t)$ without making a priori assumptions about the relationship between grant amount and sewerage capital spending, as well as the relationship between compliance and sewerage capital spending. That being said, Callaway and Sant'Anna (2020a) is much more careful about aggregating these results together and conducting inference,[55] which is why we rely on Callaway and Sant'Anna (2020a) for our main results.

A stacked estimator can be more complicated than Callaway and Sant'Anna (2020a), so in this section, we discuss how we construct stacks, as well as inference and identification with a stacked estimator.

Each stack S is defined by a single municipality, called municipality s, which is labeled as treated in stack S. Stack S also includes municipalities that receive grants at least seven years after the treated municipality, which are labeled as controls. Municipalities that receive grants early (specifically, those treated before 1979) are only in stacks in which they are the treated unit. Municipalities that receive grants too late to have any controls of their own (those treated after 1981) are only included as controls in other stacks. Municipalities that receive grants in the middle of our study period (from 1979 to 1981) can be a treated unit in one stack, while observations of that municipality from before it received a CWA grant can be controls in other stacks.

We can estimate the reduced form effect of grant receipt on sewerage capital spending for municipality s, denoted β_s^{2X2}, by comparing spending in municipality s to spending in the control municipalities in stack S before and after t_s^*, the year in which municipality s becomes treated. For a concrete example, consider a municipality treated in 1972. In the stack defined by that municipality, we can estimate a simple two group difference-in-difference where treatment turns on for the treated group in 1972. This compares sewerage capital spending in the treated municipality to spending in the control municipalities before and after 1972. The control group consists of every municipality treated in 1979 or later. This stack will include observations of both the treated and control units in each year from 1956 to 1978. Since there are no observations from 1979 or later in this stack, treatment never turns on for the control group.

Using units treated at least seven years after municipality s as controls is not an arbitrary choice. As shown in Figures 3.3 and 3.4, both grant amount and sewerage capital spending return to near pre-treatment levels by 6 years after treatment. If we only used units treated more than seven years in the future as controls, we could look at a longer post period, but our control group would be very small. We could have a larger control group if we used a shorter post period, but the resulting estimates would not capture increases in

[55] Callaway and Sant'Anna (2020a) is also more careful about controlling for confounders, though our analysis does not have any controls.

sewerage capital spending caused by CWA grants that took several years to appear in the data. Using units treated seven years in the future as controls is a compromise between these two options that likely captures the majority of the effects of CWA grants.

Our estimates of the β_s^{2X2} are noisy, so instead of interpreting individual β_s^{2X2}, we create summary measures of the effect of grant receipt on sewerage capital spending by aggregating the β_s^{2X2} together. Before doing so, we examine heterogeneity in pass-through by both grant size and compliance in Figures 3.5 and 3.6. On the bias/variance trade-off, these figures are very low bias, and we can use them to determine the best way to obtain lower variance summary measures of pass-through. The shapes of these figures motivate us to separately summarize pass-through for compliant municipalities, non-compliant municipalities that received small grants and non-compliant municipalities that received large grants.

We summarize the effect of grants on sewerage capital spending for each of these groups with $\beta_{rf}^{stacked}$ from equation 3.6. In practice, $\beta_{rf}^{stacked}$ is a weighted average of the β_s^{2X2} from each stack. The weights on each β_s^{2X2} in our estimate of $\beta_{rf}^{stacked}$ are based on stack size and within-stack variance of D_{it} (Goodman-Bacon, 2021a). We construct our stacks in such a way that each β_s^{2X2} is weighted equally by our regression. To do this, we collapse our control observations in each stack to yearly means. In the collapsed data, each stack contains two observations for each year (one treated observation and one control observation) for 16 years of pre-treatment data and 7 years of post-treatment data. Since each stack has the same number of observations and the treated unit in each stack is treated for the same percent of the time, each stack has the same weight.

We cannot construct standard errors clustered at the municipality level using this collapsed data. Instead, we construct pair-bootstrap standard errors. To do this, we take a random sample (with replacement) of observations in the unstacked data, clustering at the municipality level to preserve any dependence of error terms across time. We then re-form our stacks and perform our estimates. We repeat this process 1000 times and use the results to calculate bootstrap standard errors with equation 3.18.

In the main text, we assume that treatment timing is uncorrelated with trends in local spending, but we can state the assumption needed to interpret our stacked estimates as the ATT more precisely. Remember that the coefficient from our stacked estimator is an unweighted average of the difference in sewerage capital spending between treated and control municipalities in each stack before and after treatment. Estimating this simple difference-in-difference in a given stack S (defined by municipality s) yields β_s^{2X2}, which we can express in terms of means as

$$\beta_s^{2X2} = (\bar{y}_s^{post} - \bar{y}_s^{pre}) - (\bar{y}_{\sim s}^{post} - \bar{y}_{\sim s}^{pre})$$

This is the difference in the outcome variable in municipality s before and after t_s^* minus the difference in the outcome variable in the control municipalities in stack S before and after t_s^*. If we define $y_{it}(1)$ as the potential outcome if municipality i is treated in year t, and $y_{it}(0)$ as the potential outcome if municipality i is not treated in year t, then we can express β_s^{2X2} in terms of potential outcomes as

$$\begin{aligned}\beta_s^{2X2} =&(E[y_{it}(1)|c=s,post]-E[y_{it}(0)|i=s,pre]) \\ &-(E[y_{it}(0)|i\neq s,post]-E[y_{it}(1)|i\neq s,pre]) \\ =&E[y_{it}(1)-y_{it}(0)|i=s,post]+E[\Delta y_i(0)|i=s]-E[\Delta y_i(0)|i\neq s] \\ =&ATT_s+E[\Delta y_i(0)|i=s]-E[\Delta y_i(0)|i\neq s]\end{aligned}$$

ATT_s is the effect of treatment in municipality s, so β_s^{2X2} consists of a causal parameter and an identifying assumption. In this case, to interpret β_s^{2X2} as the effect of treatment in municipality s, we must assume that, on average, municipality s would have followed parallel trends with the control municipalities in stack S in absence of treatment.

Since $\beta^{stacked}$ (the result from a stacked difference in difference) is an unweighted average of the β_s^{2X2} from each stack, we can express $\beta^{stacked}$ as

$$\begin{aligned}\beta^{stacked} &= \frac{1}{S}\sum_{s=1}^{S}(ATT_s+E[\Delta y_i(0)|i=s]-E[\Delta y_i(0)|i\neq s]) \\ &= \frac{1}{S}\sum_{s=1}^{S}ATT_s+\frac{1}{S}\sum_{s=1}^{S}(E[\Delta y_i(0)|i=s]-E[\Delta y_i(0)|i\neq s])\end{aligned}$$

so our identifying assumption is that $\sum_{s=1}^{S}(E[\Delta y_i(0)|i=s]-E[\Delta y_i(0)|i\neq s])=0$. Note that this does not require us to assume that parallel trends holds in *every* stack, only that any violations of parallel trends cancel out when we take an average of the effect in each stack.

3.7.4 Data Details

3.7.4.1 Grant Data

We begin with grant data from the EPA's Grant Information Control System, which we obtained through a Freedom of Information Act request. This data contains information on the year that the EPA distributed each grant, which municipality received the grant, the specific wastewater treatment facility the grant was designated for and the amount distributed. Keiser and Shapiro (2019) uses the same data, and Appendix Section B.4 of Keiser and Shapiro (2019) demonstrates its accuracy.

Some grants are explicitly for construction (and include a plant code), while others do not have a plant code. It is unclear to what extent these grants were precisely for upgrading wastewater treatment plants, so we drop grants that did not have a specific facility code. This restricts our sample to 33,429 grants. We also drop grant records that are missing the year in which they were distributed, which drops another 3533 observations, as well as the 475 grants distributed to municipalities outside the contiguous U.S. Some municipalities received grants in 1960-72 through predecessor programs similar to the CWA. At least some of these precursor grants are included in our data from the EPA, but we did not assess the quality or completeness of information of grants from before 1972, so we exclude the 5248 observations of precursor grants from our analysis. This leaves us with records of 25,997 grants to 12,291 facilities.

If municipalities in our sample received any of the grants that we drop, our pass-through results could be biased upwards in two ways. First, when units act as a treated group, any increases in spending that these grants induce will appear in the reduced form estimate of the effect of grant receipt on targeted spending, but the the associated first stage estimates will not capture increases in grant funding; second, the treatment effects of these grants will change over time (in particular, spending will return to pre-treatment levels as grant funding runs out), and when units act as controls, these changes in treatment effects will be subtracted out of our reduced form estimates.

3.7.4.2 Clean Watershed Needs Survey Data

We define whether a facility was in compliance with the CWA's capital mandate using the 1972 Clean Watershed Needs Survey, which we merge to our grant data with a unique facility code. The CWNS is an assessment of the capital investment that publicly-owned wastewater treatment facilities needed to make in order to come into compliance with the Clean Water Act and contains information on which community the facility serves, the total wastewater flowing through the facility, the treatment technology currently in place, whether the facility needs to meet standards higher than the EPA's secondary treatment requirement and whether they are currently in compliance with these requirements. This data comes from the EPA's CWNS team, and is the same data that Jerch (2018) uses to define compliance with the CWA's capital mandate.

We use a facility's answer to Question 21 on the CWNS questionnaire (reproduced in Figure 3.33) to define compliance. Question 21b asks if a facility needs to meet treatment technology requirements that are more stringent than the EPA's secondary treatment requirement. Question 21c then asks whether a facility is currently in compliance with both the EPAs secondary treatment mandate and any higher mandates. We define facilities that answered "yes" on question 21c as "compliant", and those that answer "no" as "non-compliant". Table 3.35 shows cross-tabs of these variables. Our sample includes facilities that were in compliance with the CWA's capital mandate when the CWA came into effect but were not in compliance with more stringent state standards. These facilities still faced incentives to spend CWA grants on sewerage capital, so we classify them as non-compliant.

Note that many facilities installed tertiary treatment after the CWA came into effect (USEPA, 2000). This increase was likely driven by municipalities bound by state standards or compelled by lawsuits to make upgrades beyond secondary treatment. Since we define these municipalities as non-compliant and we document an increase in sewerage capital spending in response to grant receipt in non-compliant municipalities, it is possible that municipalities used CWA grants to upgrade to tertiary treatment.

Our results are robust to dropping municipalities with facilities that were in compliance with the CWA's capital mandate but were not in compliance with more stringent state standards. Panels A and B of Table 3.37 re-estimate the results in Panels A and B of Table 3.4 after dropping non-compliant municipalities that had to satisfy state wastewater capital requirements greater than the EPA's secondary treatment requirement. There is still evidence of full pass-through in municipalities that received grants smaller than \$125 per capita and substantial crowd-out in municipalities that received grants larger than this amount.

We re-estimate the point where costs were no longer increasing in grant size using this sample in Table 3.38. This produces an estimate that is similar to the results in Table 3.3. We re-estimate pass-through for non-compliant facilities that were not required to satisfy state wastewater capital requirements greater than the EPA's secondary treatment requirement that received grants above and below this point in panels C and D of Table 3.37. Again, the results are similar to those in Table 3.4.

Since our data on compliance is nearly 50 years old, we assessed its quality by comparing it to information on treatment technology from Table 2.4 of USEPA (2000). This table shows that, of 19,355 publicly owned wastewater treatment facilities surveyed, 9,887 were using secondary treatment or greater in 1972. This represents 51% of facilities. While we only observe whether a facility was in compliance with both EPA and state standards as opposed to just the secondary treatment technology standard, we can observe the percent of facilities that had at least secondary treatment among facilities that were only required to meet the EPA's standard. Column 1 of Table 3.35 shows that 56% of these facilities were using secondary treatment or greater in 1972. Given the different sample, this is relatively close to the 51% of facilities using secondary treatment or greater in 1972 in USEPA (2000).

Our definition of compliance does not directly enter into our pass-through equations. Instead, we use compliance to define sub-groups that we expect to find heterogeneous treatment effects for. If we mis-classify non-compliant municipalities as compliant, our estimates of pass-through for non-compliant municipalities will not be biased, but our estimates of pass-through for compliant municipalities would be biased upwards since they would include the effect of grant dollars that were bound by the CWA's capital mandate, and thus theoretically more likely to be spent on sewerage capital. Conversely, if we mis-classify compliant municipalities as non-compliant, our estimates of pass-through for compliant municipalities will not be biased, but our estimates of pass-through for non-compliant municipalities would be biased downwards since they would include the effect of more grant dollars that were not bound by the CWA's capital mandate.

Our definition of compliance does directly enter into our estimates of the effect of grant receipt on water revenue in Section 3.5. Mis-classifying municipalities as compliant or non-compliant would attenuate these results.

3.7.4.3 Estimating Costs

We define compliance in such a way that compliant municipalities never face any mandated costs, but compliance is not the only source of heterogeneity in mandated costs; there is also variation in mandated costs within non-compliant municipalities. The costs of upgrading depend largely on facility size (USEPA, 1973), and we use engineering estimates of the costs of upgrading from USEPA (1973) to approximate these costs.

USEPA (1973) uses data from the observed costs of upgrading plants to secondary treatment and fits this data to the number of gallons of wastewater flowing through a facility. This yields the following equation:

$$TotalCost = exp(.13732 + .77872 * ln(TotalFlow))$$

We use this equation to calculate mandated costs for each facility, then sum these costs across the facilities operated by a municipality to approximate the total costs that each municipality faces. While this measure of mandated costs is noisy, since it is based on an engineering formula, it is unlikely to be correlated with local preferences.[56]

These cost estimates never enter into any of our equations directly. Instead, we use them to determine if grants are larger or smaller than mandated costs. Using these cost estimates, we find that, while grants are similar to costs for municipalities that receive up to around $125 per capita in total grant aid, grants to municipalities that receive more than this amount are substantially larger than the estimated cost of upgrading. We then separately estimate pass-through for municipalities that received grants above and below $125 per capita in total grant aid. Noise in our cost estimates will cause us to mis-classify non-compliant municipalities between these two groups. This will attenuate our pass-trough estimates for non-complaint municipalities that received small grants, since it will cause these estimates to include non-compliant municipalities that received grants that included funding that was not bound by the capital mandate, and will bias our pass-trough estimates for non-complaint municipalities that received large grants upwards, since it will include municipalities whose grants were completely bound by the capital mandate. Noise in our cost estimates will also cause our pass-through estimates for both groups to lose precision.

[56]If we used an assessment conducted by municipalities instead, we might be concerned that self-reported total costs are correlated with preferences for spending.

3.7.4.4 Merging Facility Data to Municipal Spending Data

The CWNS contains data on the authority that operates each facility, and we use this information to merge our facility-level data to municipality-level spending data from the Census Bureau's Historical Database on Individual Government Finances. The CWNS data does not contain a unique identifier of municipalities that is consistent with our local government finance data, but since we only have a balanced panel of spending data for 216 municipalities, it is relatively straightforward to use the state and municipality name in both detests to match our municipality level data to the facilities that they operated.

3.7.4.5 Measuring Spending in Real Per Capita Dollars

We follow the previous pass-through literature by measuring grant amount and all spending in per capita terms. (Strumpf, 1998; Card and Payne, 2002; Vegh and Vuletin, 2015). All of our variables of interest are largely influenced by population, but dividing these variables by population ensures that our estimates are not biased by population growth. Figure 3.34 re-estimates the results in Figure 3.4 with population as the dependent variable, and shows that population did not change after grant receipt.

We deflate grant amount and all spending variables to 1973 dollars with the Consumer Price Index. As shown in Figures 3.35, 3.36 and 3.37, and Tables 3.39, 3.40 and 3.41, our main results are robust to measuring spending and grant amount in nominal terms.

3.7.5 Additional Figures and Tables

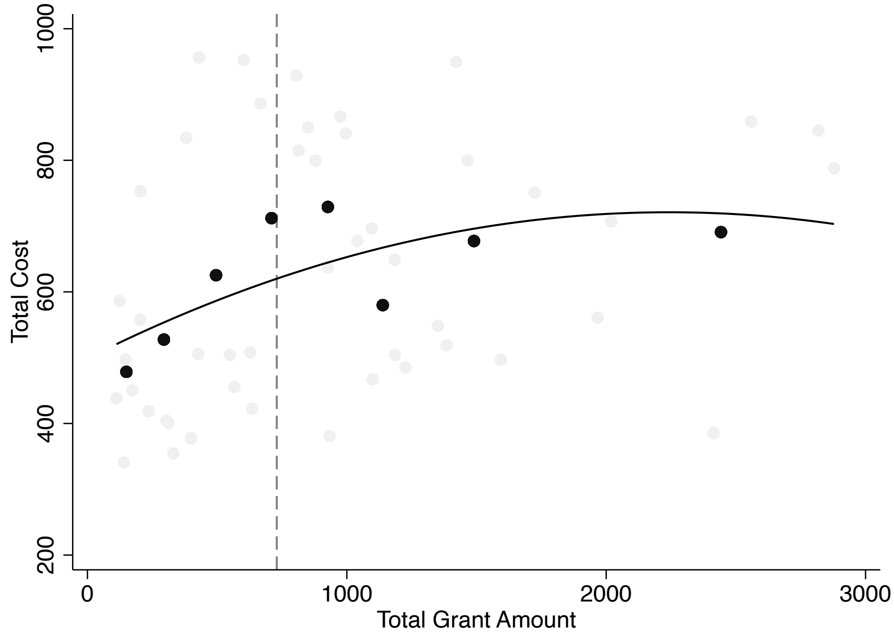


Figure 3.8: Cost Data with Quadratic Fit

Notes: This figure re-creates Figure 3.1 a quadratic fit.

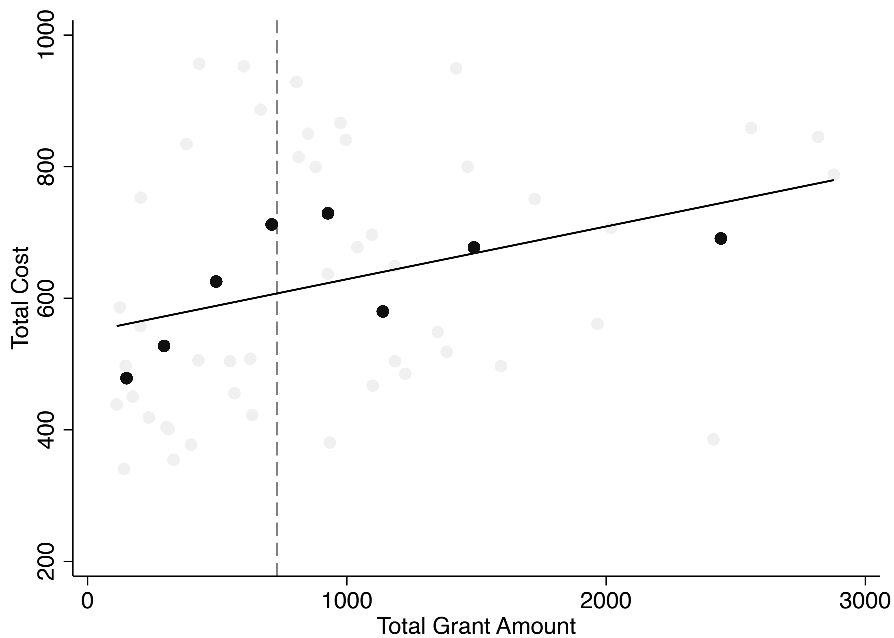


Figure 3.9: Cost Data with Linear Fit

Notes: This figure re-creates Figure 3.1 with a a linear fit.

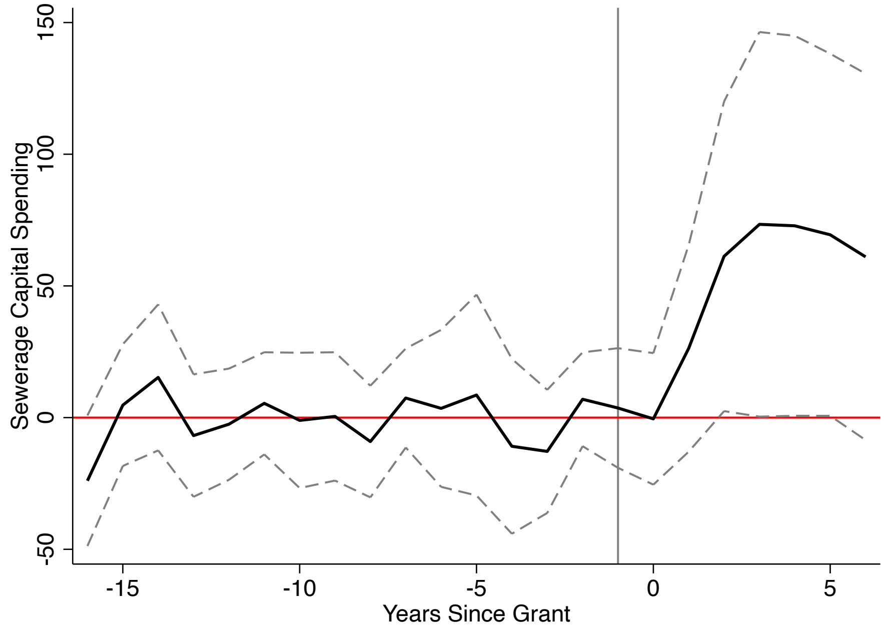


Figure 3.10: Balanced Sewerage Capital Event Study

Notes: This figure re-estimates the results in Figure 3.4 on a sample of municipalities that we observe for at least seven post-treatment periods.

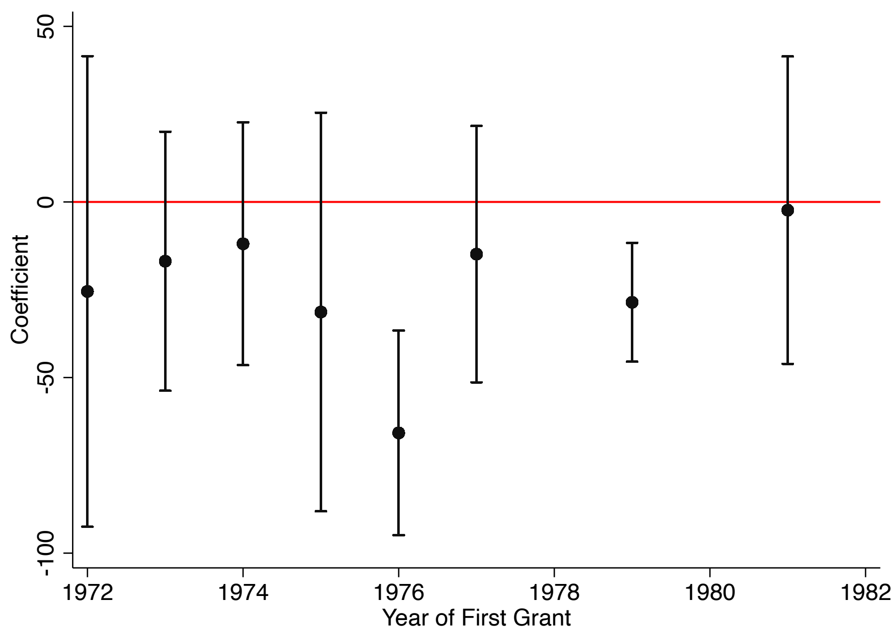


Figure 3.11: Decomposition of Main Water Revenue Estimate

Notes: This figure presents estimates of the θ_g^{2X2} from estimating $R_{it} = \alpha_0 + \theta_g^{2X2} 1\{t \geq g\} * compliant_i + \delta 1\{t \geq g\} + \omega compliant_i + \varepsilon_{it}$ on observations from each timing group g along with the 95% confidence interval of each θ_g^{2X2}. We plot each θ_g^{2X2} against the year timing group g became treated.

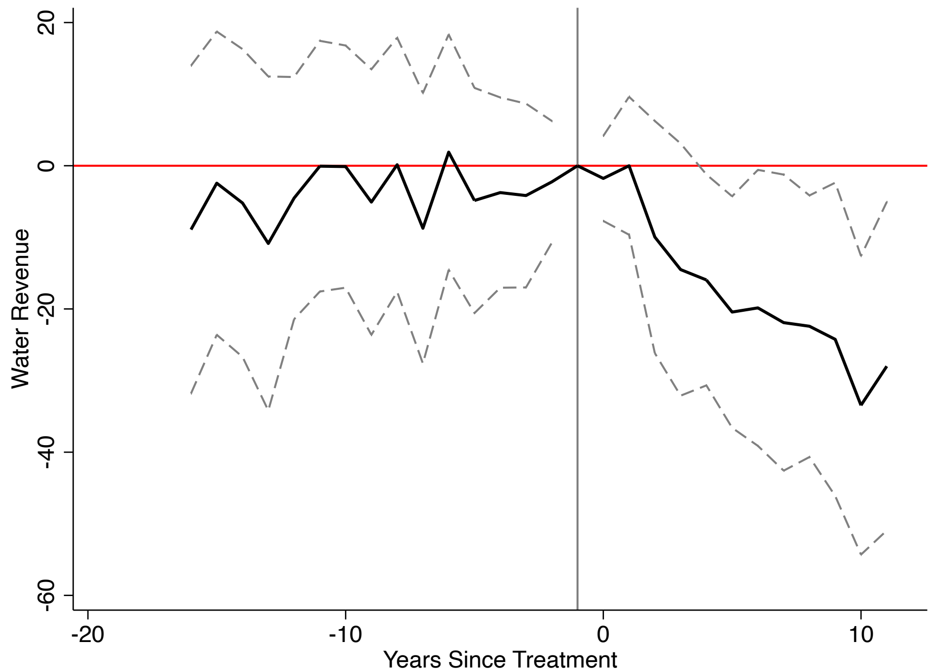


Figure 3.12: Water Revenue Event Study with Region-by-Year Fixed Effects

Notes: This figure plots the π_y and γ_y from estimating $R_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * compliant_i + \sum_{y=0}^{11} \gamma_y 1\{t - t_i^* = y\} * compliant_i + \alpha_i + \alpha_{gt} + \alpha_{rt} + \varepsilon_{it}$ with per water revenue as the dependent variable. $compliant_i$ equals one for observations of compliant municipalities. We include municipality, timing-group-by-year, and region-by-year fixed effects, α_i, α_{gt}, and α_{rt}, where g indexes the year in which a municipality received its first CWA grant and r indexes region.

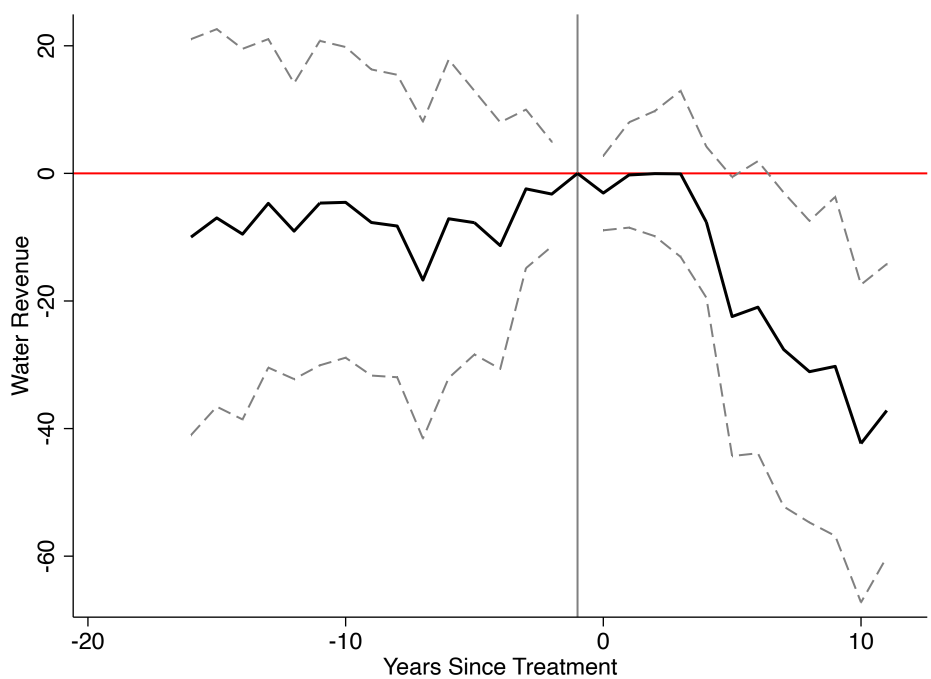


Figure 3.13: TWFE Water Revenue Event Study

Notes: This figure plots the π_y and γ_y from estimating $R_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * compliant_i + \sum_{y=0}^{11} \gamma_y 1\{t - t_i^* = y\} * compliant_i + \alpha_i + \alpha_t + \varepsilon_{it}$. α_i and α_t are municipality and year fixed effects.

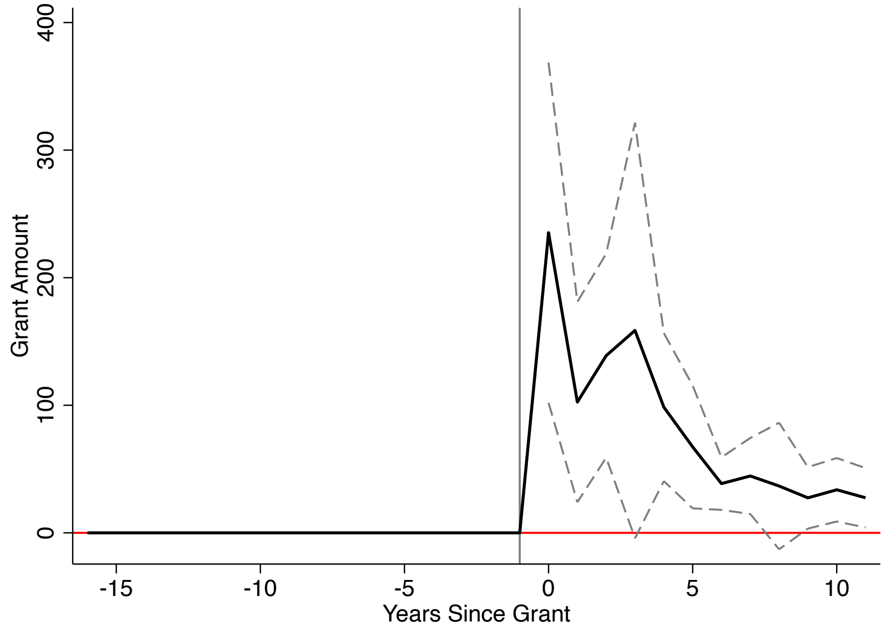


Figure 3.14: Grant Amount Event Study Using Never-Treated Controls

Notes: This figure re-creates Figure 3.3 using never-treated units as a control group.

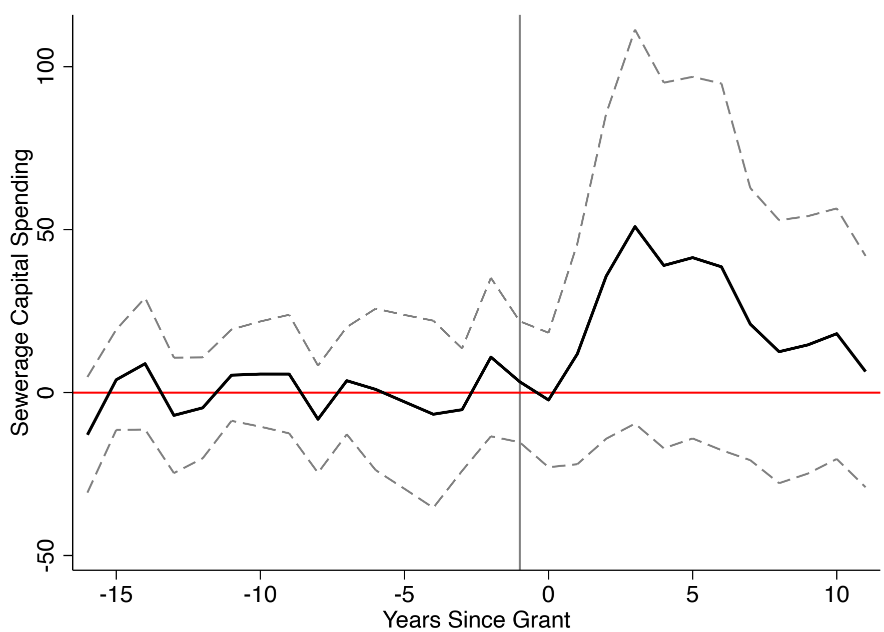


Figure 3.15: Sewerage Capital Spending Event Study Using Never-Treated Controls

Notes: This figure re-creates Figure 3.4 using never-treated units as a control group.

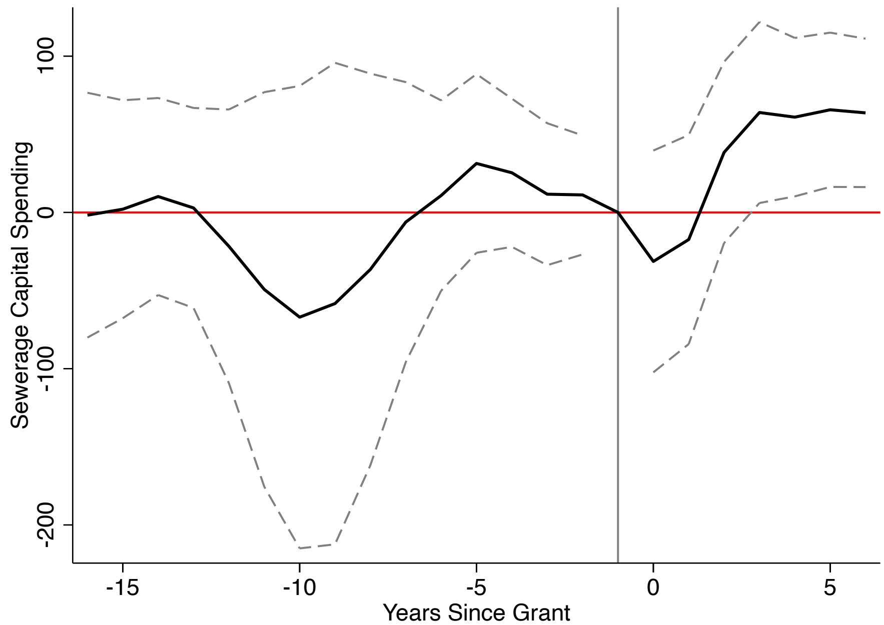


Figure 3.16: Stacked Sewerage Capital Event Study

Notes: This figure plots the π_y and γ_y from estimating $C_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * treat_{is} + \sum_{y=0}^{6} \gamma_y 1\{t - t_i^* = y\} * treat_{is} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its}$. i indexes municipalities, t indexes years and s indexes stacks. $treat_{is}$ equals one for observations of the treated municipality in each stack after that municipality received its first CWA grant. The dependent variable is per capita sewerage capital spending and α_{is} and α_{ts} are stack by municipality and stack by year fixed effects.

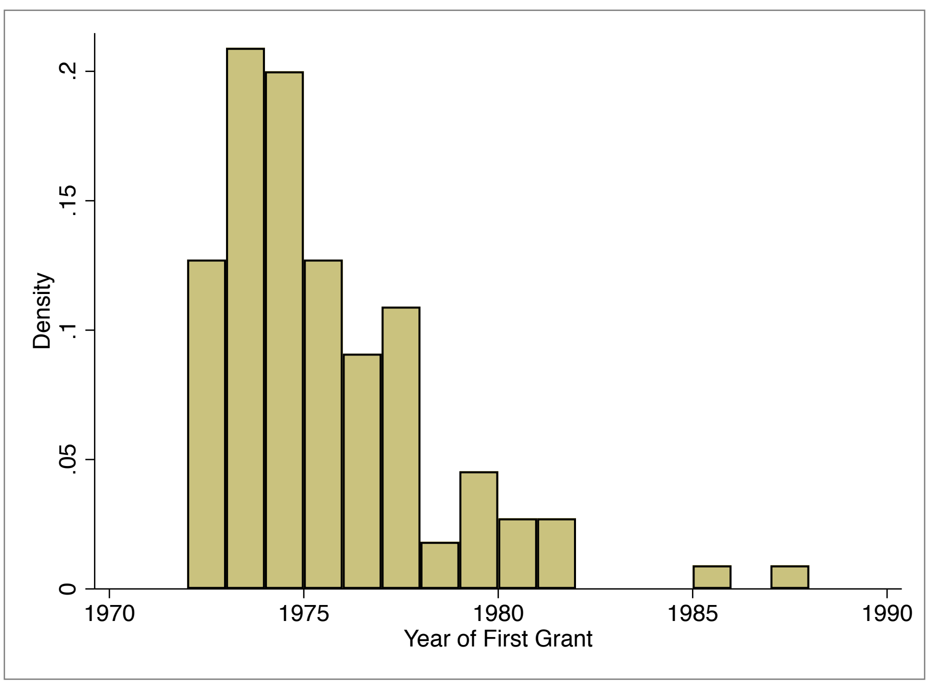


Figure 3.17: Timing Groups

Notes: This figure shows how many municipalities received their first CWA grant in each year from 1972 to 1988.

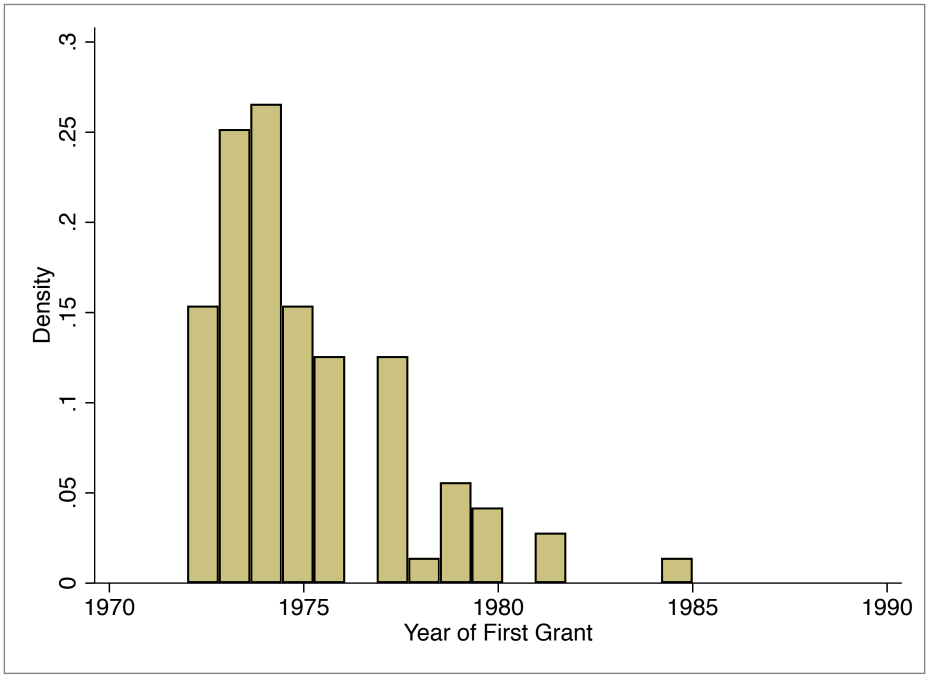


Figure 3.18: Non-Compliant

Notes: This figure re-creates Figure 3.17 for non-compliant municipalities.

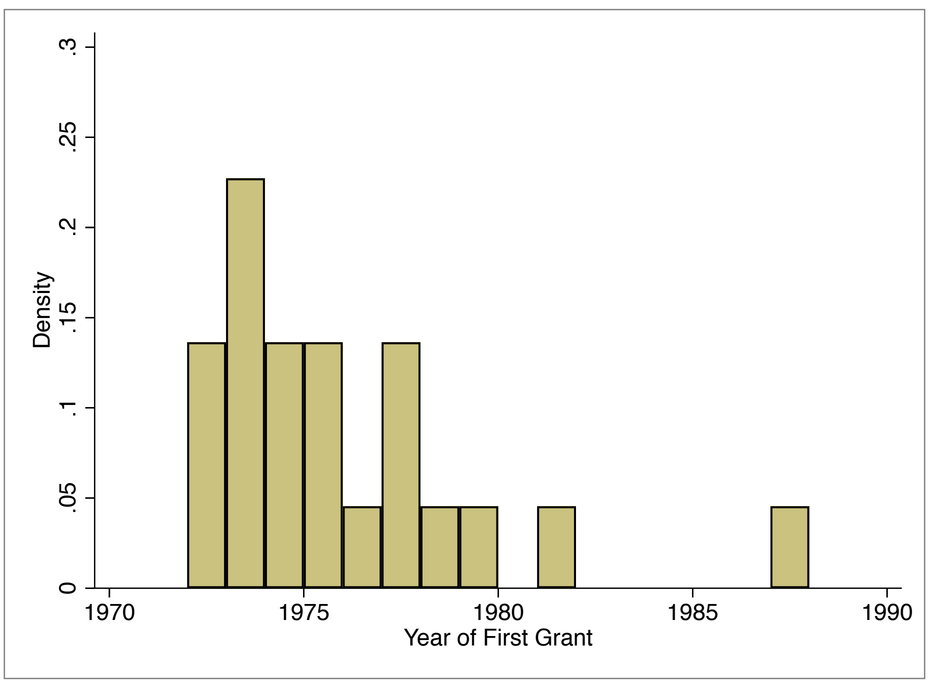


Figure 3.19: Timing Groups for Compliant Municipalities

Notes: This figure re-creates Figure 3.17 for compliant municipalities.

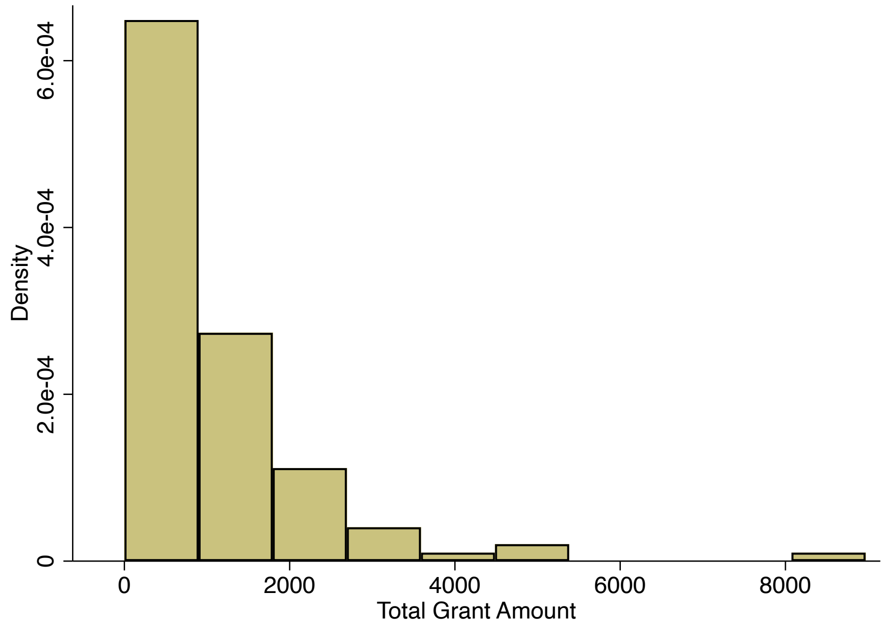


Figure 3.20: Distribution of Grant Size

Notes: This figure shows the distribution of total grant dollars.

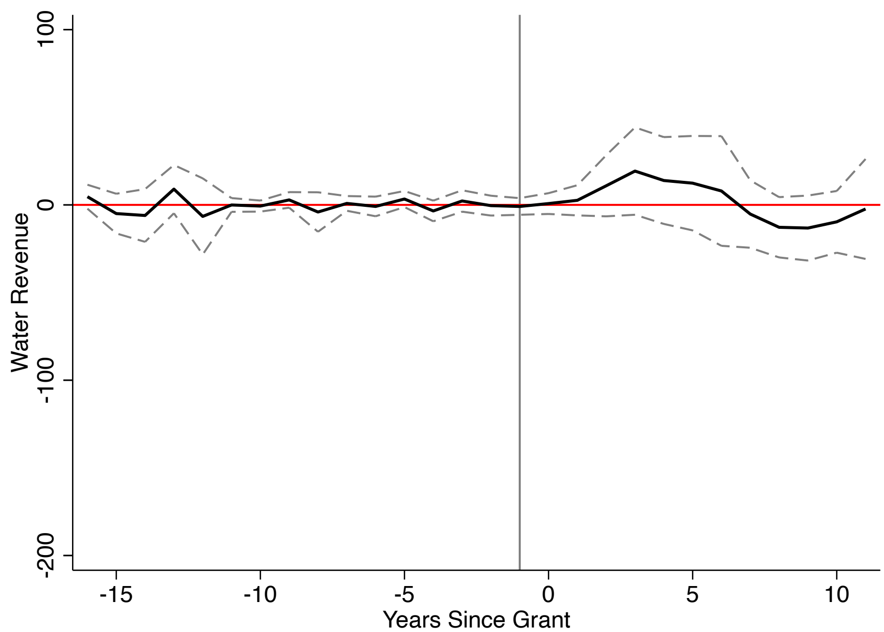


Figure 3.21: Water Revenue Did Not Change After Grant Receipt in Non-Compliant Municipalities

Notes: This figure presents $\widehat{\theta}_D(e)$ for each year e relative to treatment, where $\widehat{\theta}_D(e) = \sum_{g=1972}^{1985} \sum_{t=1957}^{1987} 1\{t-g=e\}\widehat{ATT(g,t)}P(G = g|t-g=e)$ for non-compliant municipalities. Water revenue is the dependent variable.

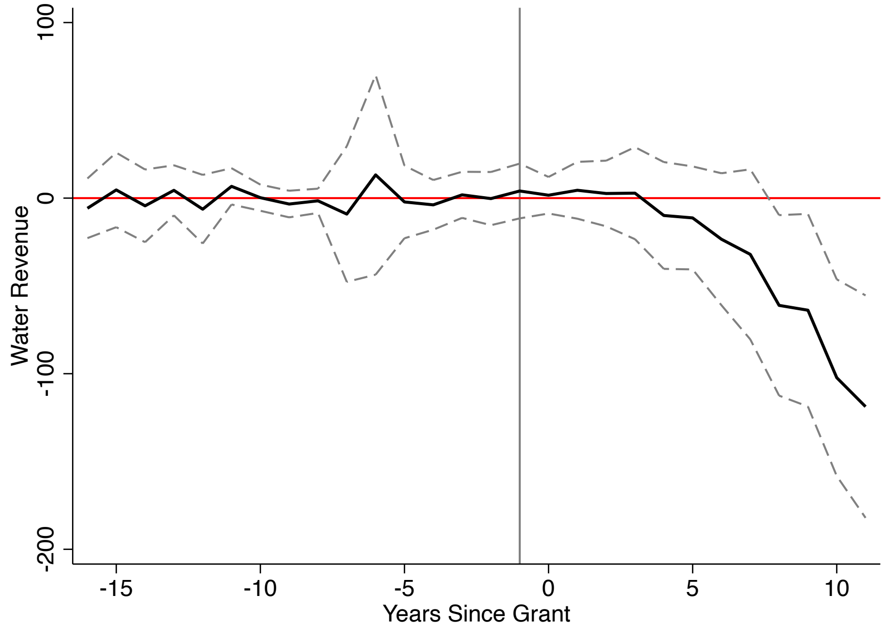


Figure 3.22: Water Revenue Decreased After Grant Receipt in Compliant Municipalities

Notes: This figure re-creates Figure 3.21 for compliant municipalities

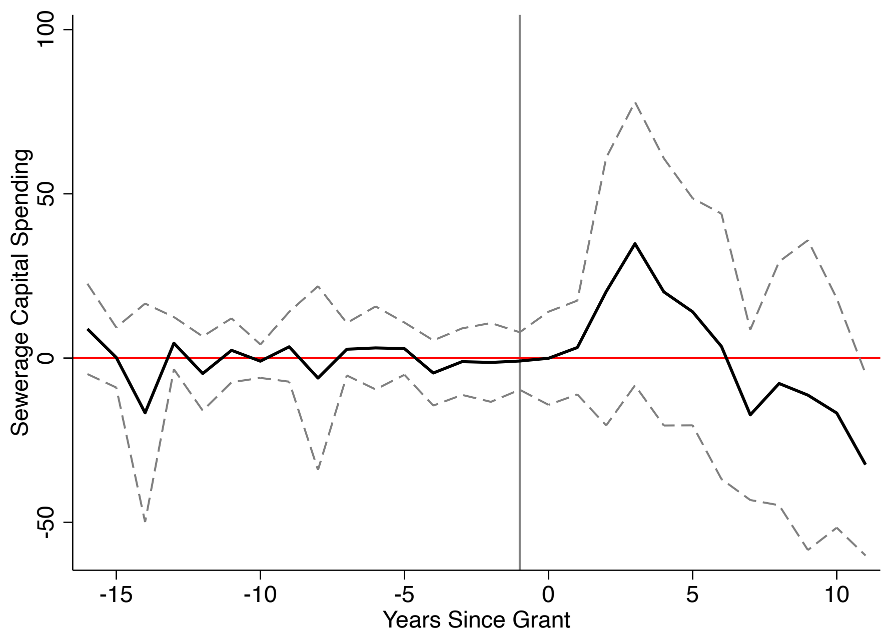


Figure 3.23: Non-Complaint Municipalities Raised Water Bills When Grants Were Too Small

Notes: This figure re-estimates the results in Figure 3.21 on sub-samples of municipalities where grants are smaller than estimated costs.

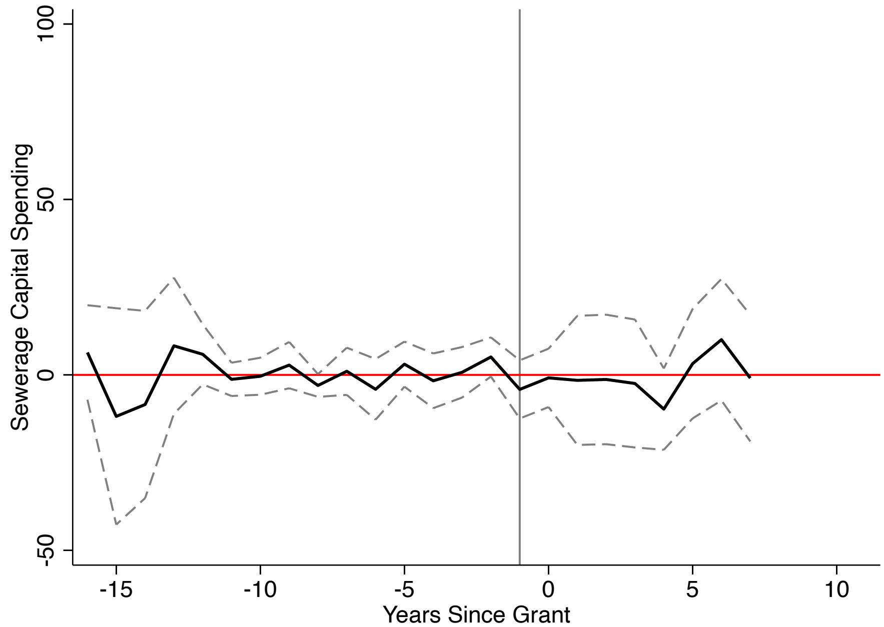


Figure 3.24: Non-Complaint Municipalities Did Not Change Water Bills in Response to Large Grants

Notes: This figure re-estimates the results in Figure 3.21 on sub-samples of municipalities where grants are larger than estimated costs.

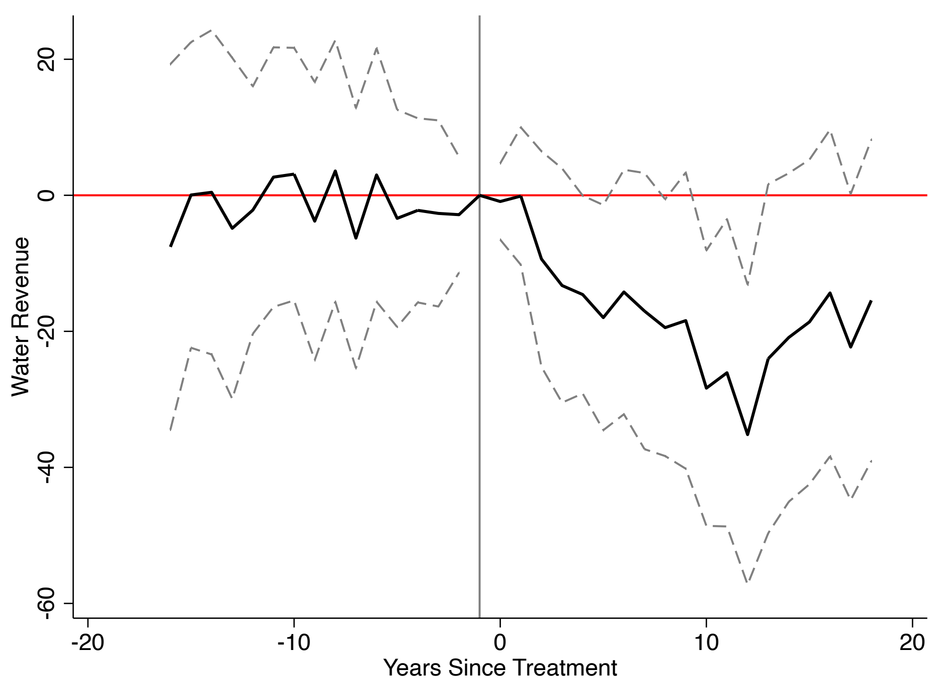


Figure 3.25: The Effect of Grants on Water Revenue Flattens Out

Notes: This figure re-estimates the results in Figure 3.7 on municipalities treated before 1982.

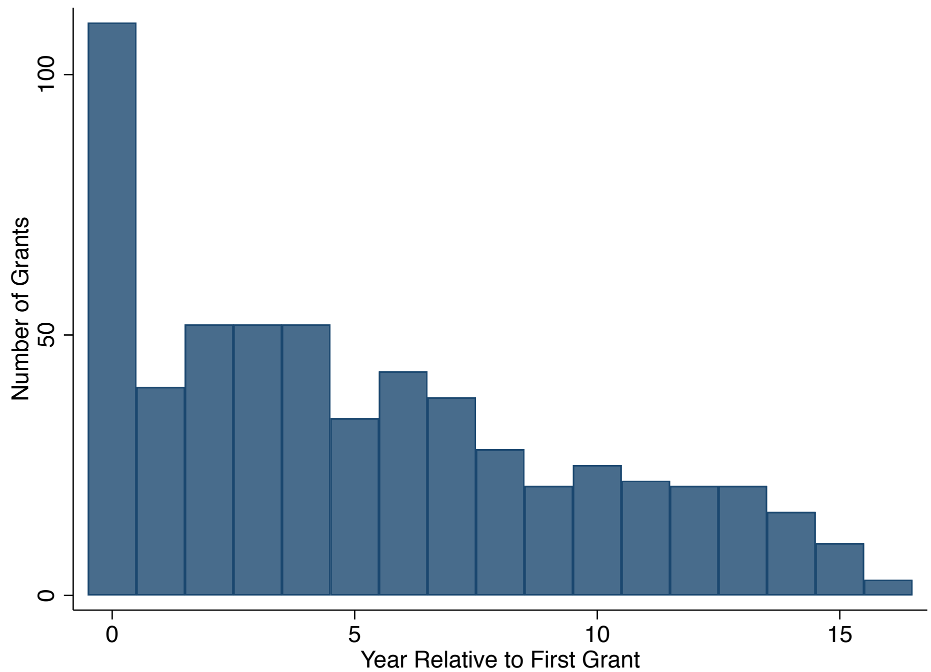


Figure 3.26: Distribution of Grants Over Time

Notes: This figure shows the number of grants that municipalities received each year relative to the year of the municipality's first grant.

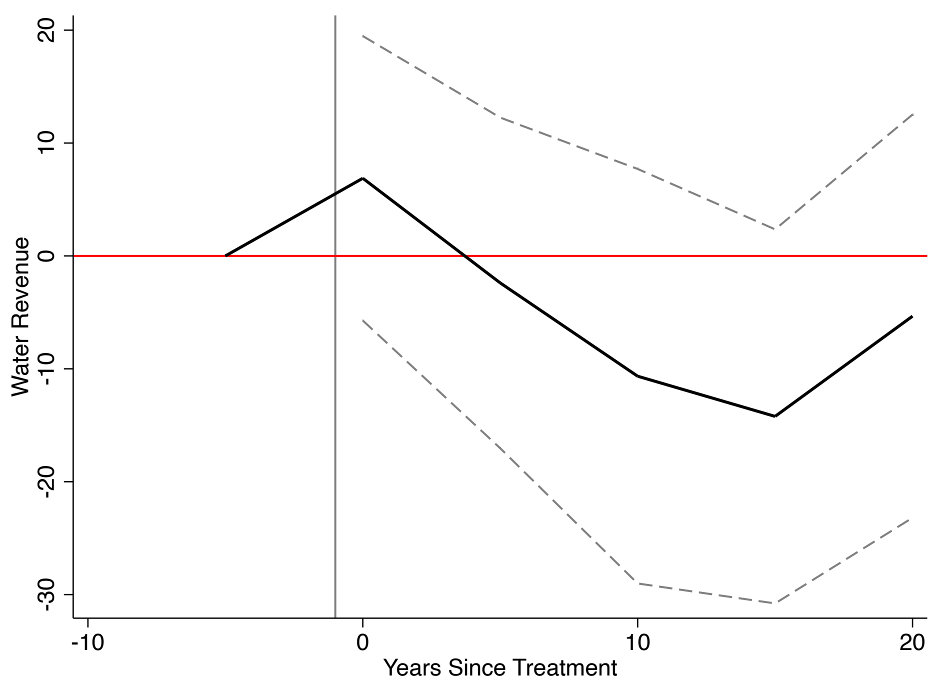


Figure 3.27: Water Revenue Event Study with Alternative Data

Notes: This figure re-estimates the results from Figure 3.7 on an alternative dataset.

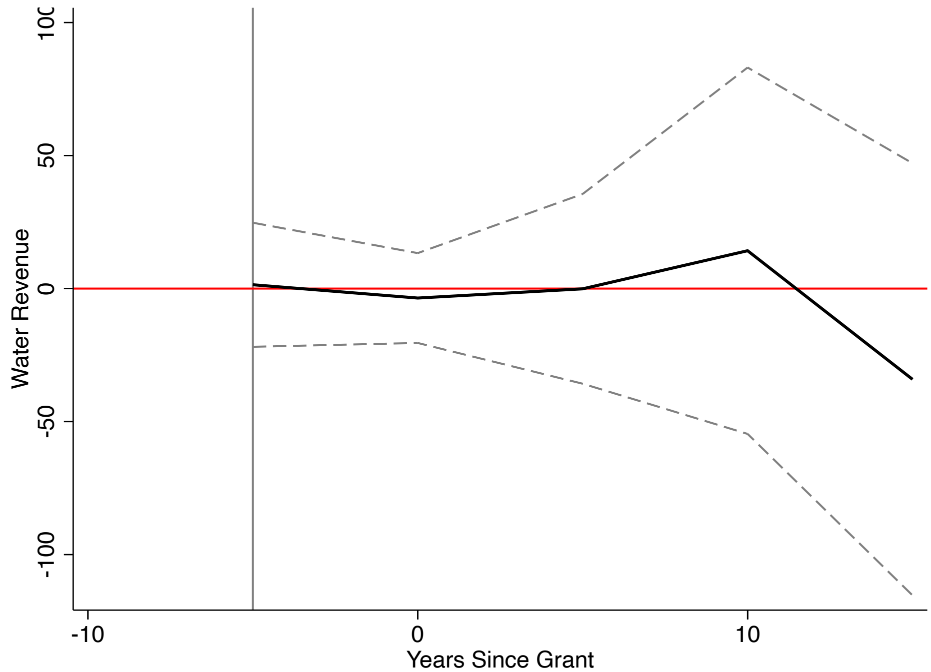


Figure 3.28: Non-Compliant Water Revenue Event Study with Alternative Data

Notes: This figure re-estimates the results from Figure 3.21 on an alternative dataset.

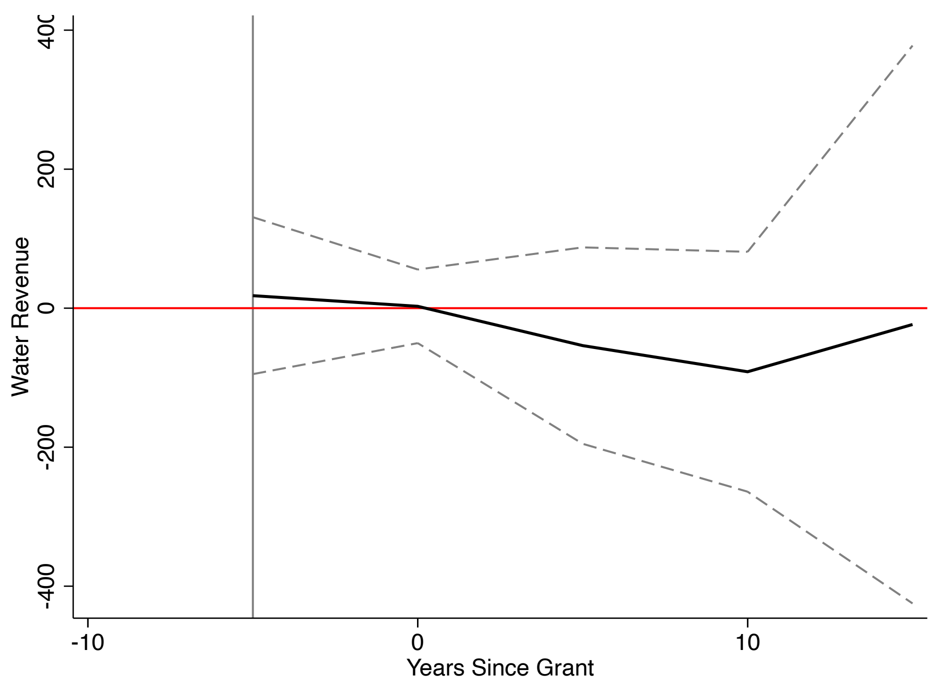


Figure 3.29: Compliant Water Revenue Event Study with Alternative Data

Notes: This figure re-estimates the results from Figure 3.22 on an alternative dataset.

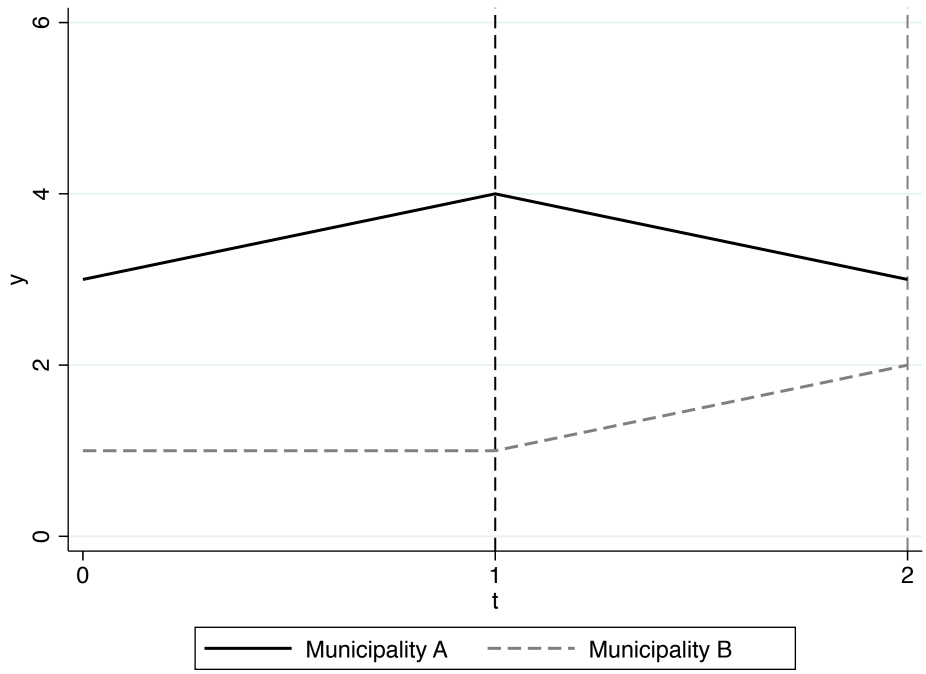


Figure 3.30: Example of Problems with Binary TWFE

Notes: This figure shows a hypothetical situations where two municipalities are treated at different times. Municipality A is treated in period 1 and Municipality B is treated in period 2.

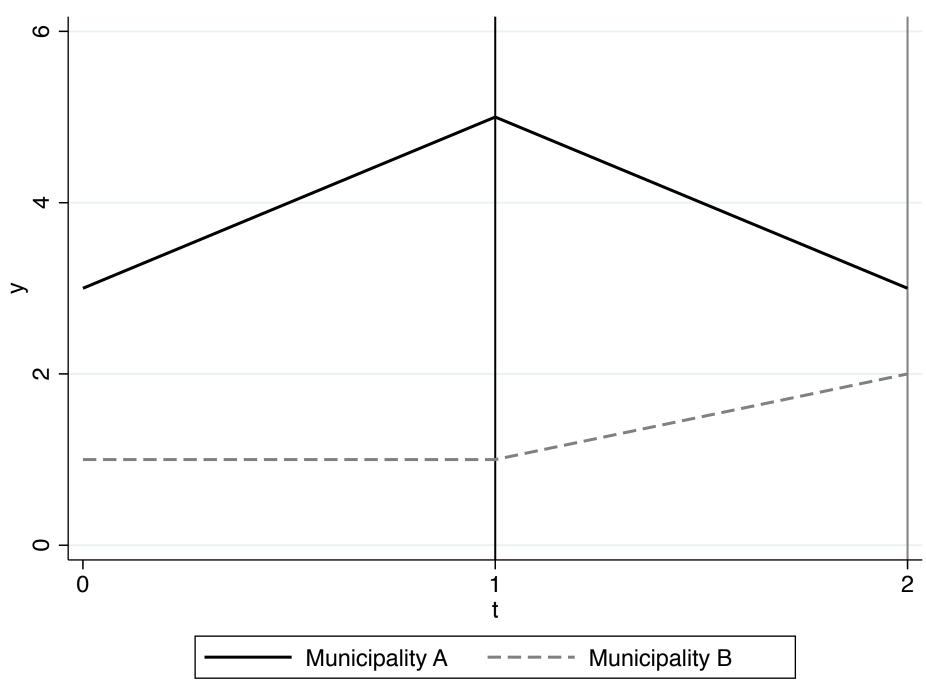


Figure 3.31: Example of Problems with Continuous TWFE

Notes: This figure shows another hypothetical situations where two municipalities are treated at different times.

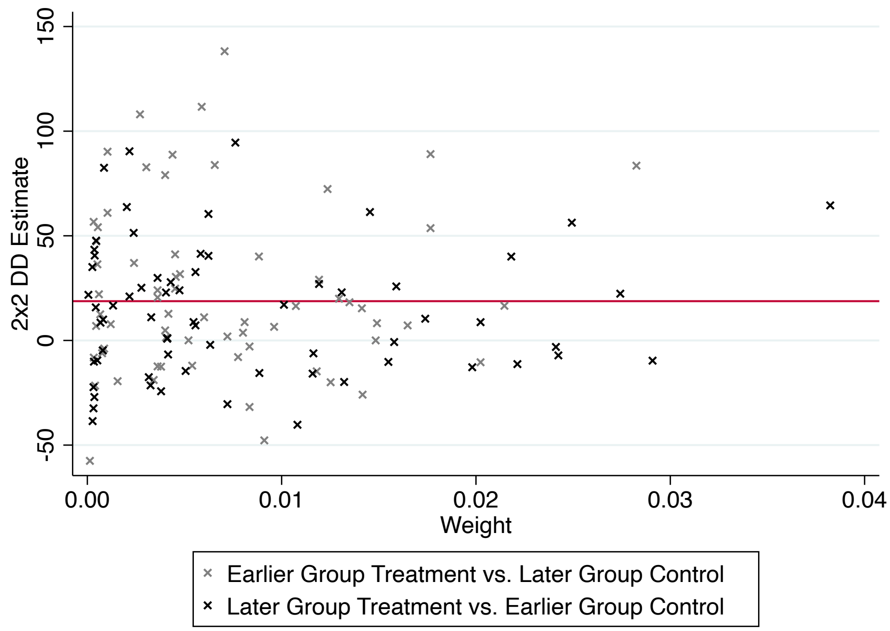


Figure 3.32: Bacon Decomposition of Sewerage Capital Spending TWFE Estimate

Notes: This figure presents a scatterplot of all possible two-group/two period DD estimators and their associated weights. These estimates and weights come from a decomposition of β_{rf}^{TWFE} from $C_{it} = \alpha_0 + \beta_{rf}^{TWFE} D_{it} + \alpha_i + \alpha_t + \varepsilon_{it}$ made using Goodman-Bacon et al. (2019). C_{it} is per capita sewerage capital expenditure, D_{it} is a dummy variable that equals one after grant receipt and α_i and α_t are municipality and year fixed effects.

Authority and Facility No. ▶ __ __ - __ __ __ __ - __ __ __

Section III – CATEGORY I – ASSESSMENT OF NEEDS BY TYPE AND COST

ASSESSMENT OF NEEDS TO ACHIEVE REQUIRED LEVEL OF SECONDARY TREATMENT	*Mark appropriate box for each item*	**EPA USE ONLY**
21a. Can this plant meet water quality standards applicable to the stream segment to which it discharges by a level of treatment LESS than defined as secondary treatment by EPA?	1 [ ] Yes 2 [ ] No 3 [ ] Not known 4 [ ] Not applicable; no discharge to waters	
b. What level of secondary treatment must the discharge from this plant meet by July 1, 1977?	1 [ ] Secondary treatment level as defined by EPA, **OR** 2 [ ] Higher level of secondary treatment required by State – *Specify* ↴ Higher level secondary treatment Nature of State action	
c. Does the discharge from this plant **NOW** meet the level of secondary treatment identified in 21b?	1 [ ] Yes 2 [ ] No	
d. Will the discharge from this plant meet on July 1, 1977, the level of secondary treatment identified in 21b? *(Give consideration to changes in flow and concentration of influents and to the changes in treatment capability now under construction or provided for in approved grants.)*	1 [ ] Yes – *SKIP to item 25* 2 [ ] No	
e. Which approaches will be used to enable this existing or proposed plant to meet the secondary treatment level identified in 21b? (1) Addition of land disposal as a means of treatment	1 [ ] Yes 2 [ ] No	
(2) New plant – no replacement of existing plant	1 [ ] Yes 2 [ ] No	
(3) Replacement plant	1 [ ] Yes 2 [ ] No	
(4) Modification – no change in capacity or treatment level	1 [ ] Yes 2 [ ] No	
(5) Modification – increase in capacity	1 [ ] Yes 2 [ ] No	
(6) Modification – increase in treatment level	1 [ ] Yes 2 [ ] No	
(7) Improved operation and maintenance, increase staffing	1 [ ] Yes 2 [ ] No	
(8) Reduce infiltration	1 [ ] Yes 2 [ ] No	

Page 15

Figure 3.33: CWNS Compliance Question

Notes: This figure reproduces the question from the 1972 CWNS that we use to define compliance.

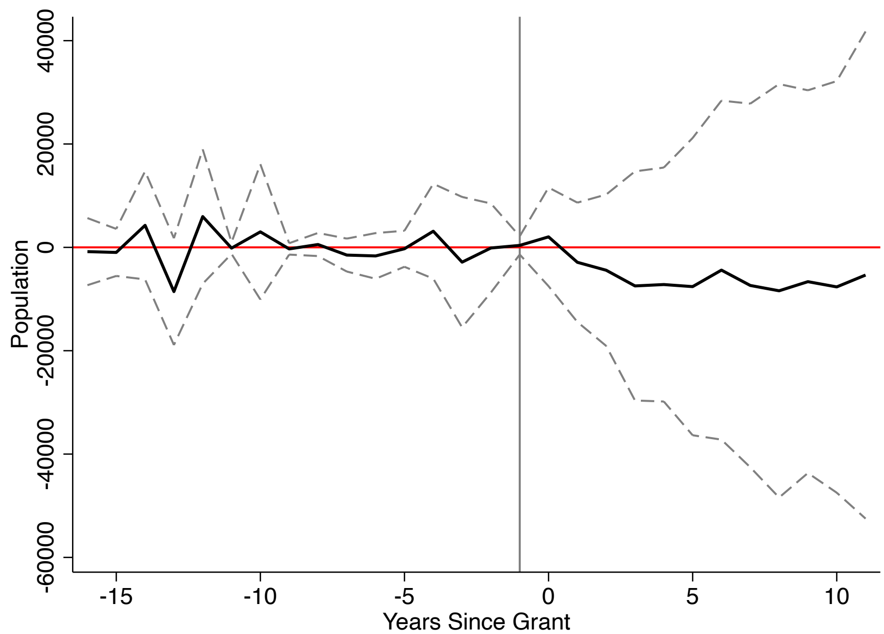


Figure 3.34: Population Did Not Change After Grant Receipt

Notes: This figure shows the reduced form relationship between grant receipt and population by presenting the $\widehat{\theta}_D(e)$ for each year e relative to treatment, where $\widehat{\theta}_D(e) = \sum_{g=1972}^{1985}\sum_{t=1957}^{1987} 1\{t-g=e\}\widehat{ATT(g,t)}P(G=g|t-g=e)$. g indexes timing groups and t indexes years. $\widehat{ATT(g,t)}$ is the difference between changes in population in units in group g in periods t and $g-1$ and changes in population in units not yet treated at time t in periods t and $g-1$.

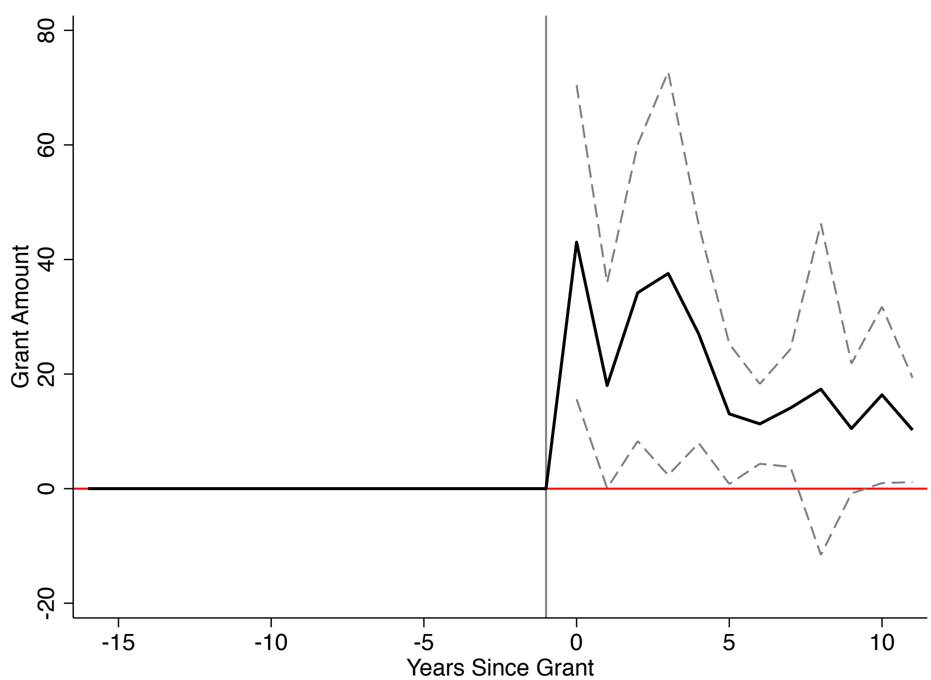


Figure 3.35: Grant Amount Event Study (Nominal Dollars)

Notes: This figure re-creates Figure 3.3 measuring grant amount in nominal dollars.

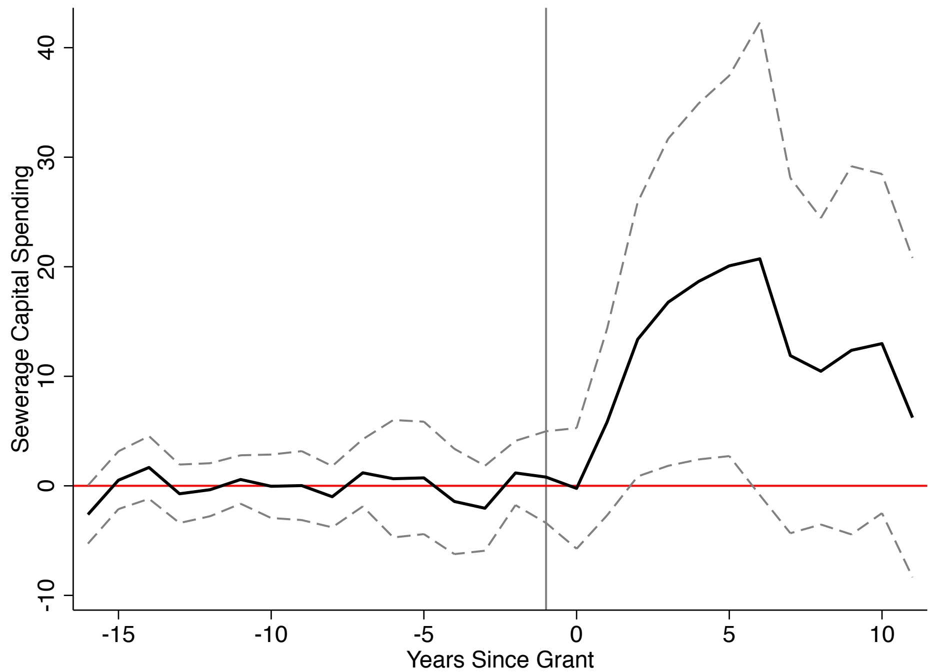


Figure 3.36: Sewerage Capital Spending Event Study (Nominal Dollars)

Notes: This figure re-creates Figure 3.4 measuring sewerage capital spending in nominal dollars

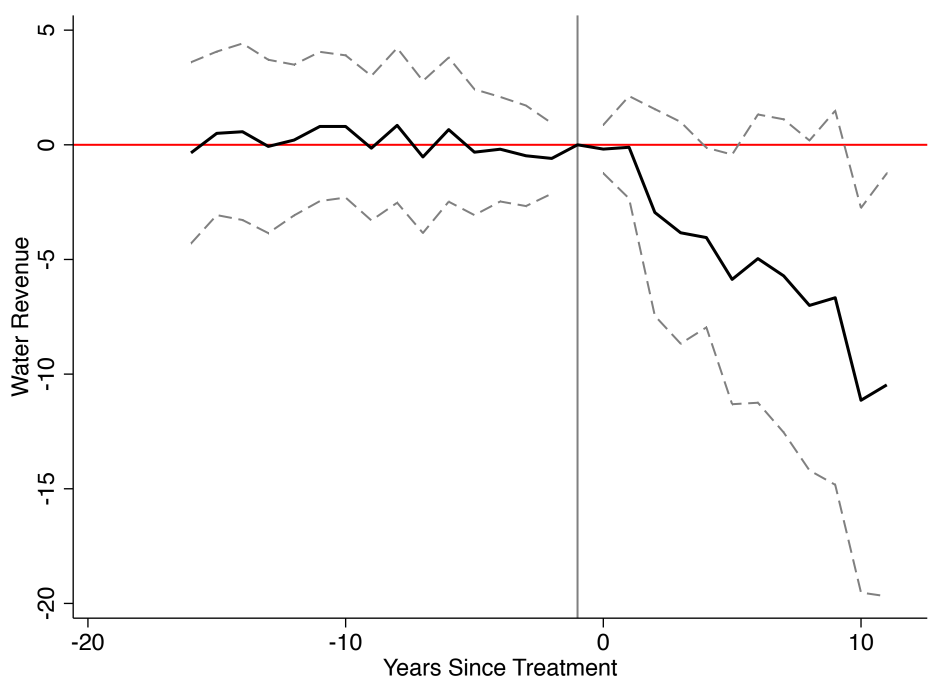


Figure 3.37: Water Revenue Event Study (Nominal Dollars)

Notes: This figure re-creates Figure 3.7 measuring water revenue in nominal dollars.

Table 3.10: 1970 Summary Statistics by Facility Compliance and Grant Size

	(1) Compliant	(2) NC & < $125	(3) NC & >= $125
Population	115,501	398,565	395,286
	(73,126)	(722,518)	(1077607)
Total Revenue	253.13	290.61	296.40
	(138.17)	(136.48)	(166.63)
Total IGR	52.84	50.67	72.72
	(62.93)	(40.28)	(82.90)
Federal IGR	11.43	9.14	11.89
	(15.64)	(8.89)	(12.89)
Nonwater Utility Rev	21.18	32.53	9.42
	(44.79)	(60.27)	(30.16)
Property Tax Revenue	82.21	99.47	97.59
	(72.75)	(94.32)	(69.14)
Water Utility Revenue	23.10	22.93	22.90
	(9.10)	(8.84)	(8.50)
Total Expenditure	277.26	302.29	311.73
	(154.56)	(145.95)	(179.20)
Total IG Exp	1.93	4.79	4.38
	(4.54)	(9.98)	(10.07)
Capital Outlays	79.35	68.87	60.84
	(57.33)	(41.85)	(34.90)
Highway Exp	21.16	21.25	22.37
	(8.38)	(9.54)	(9.38)
Sewerage Capital Outlay	8.94	6.66	5.87
	(9.61)	(6.50)	(7.01)
Sewerage Non-capital Exp	5.50	4.91	5.12
	(3.13)	(3.37)	(2.97)
Nonwater Utiltiy Exp	22.46	31.23	11.28
	(49.25)	(56.94)	(38.13)
Water IG Exp	1.98	2.91	2.07
	(1.91)	(3.22)	(2.09)
Water Non-capital Exp	13.34	12.03	13.35
	(4.74)	(7.16)	(4.99)
Water Capital Outlay	13.45	11.84	6.27
	(19.01)	(16.15)	(6.83)
Observations	22	34	56

Notes: This table presents summary statistics for municipalities with compliant facilities, non-compliant facilities that received small grants, and non-complaint facilities that received large grants in 1970. All covariates aside from the facility compliance dummy and population represent per capita values.

Table 3.11: Observable Characteristics Do Not Predict Grant Size or Compliance

	(1) Grant >= 125	(2) Compliant
ln Population	0.0315	-0.0662*
	(0.0595)	(0.0370)
Total Revenue	-0.00121	-0.00276
	(0.00299)	(0.00193)
Total IGR	0.00311*	-0.000142
	(0.00180)	(0.00140)
Federal IGR	0.00159	0.000359
	(0.00268)	(0.00194)
Nonwater Utility Rev	-0.00371	-0.00381
	(0.00288)	(0.00236)
Property Tax Revenue	-0.00127	0.000757
	(0.00118)	(0.000899)
Total Expenditure	0.00160	0.00162
	(0.00318)	(0.00211)
Total IG Exp	0.00249	-0.00349
	(0.00449)	(0.00358)
Capital Outlays	-0.00398	0.000949
	(0.00274)	(0.00207)
Highway Exp	-0.00155	0.00177
	(0.00503)	(0.00428)
Sewerage Capital Outlay	0.00267	-0.00294
	(0.00509)	(0.00351)
Sewerage Non-capital Exp	-0.00203	0.00719
	(0.0147)	(0.0133)
Nonwater Utiltiy Exp	0.00157	0.00500*
	(0.00263)	(0.00276)
Water IG Exp	-0.0128	-0.0188
	(0.0281)	(0.0230)
Water Non-capital Exp	0.0269*	-0.00377
	(0.0160)	(0.00964)
Water Capital Outlay	-0.00158	0.00812
	(0.00534)	(0.00570)
Water Utility Revenue	-0.00519	-0.00155
	(0.0102)	(0.00782)
Observations	112	112

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents results from regressing dummies indicating municipalities that received grants totaling more than $125 (in column 1) and compliant municipalities (in column 2) on potential predictive covariates from 1971. All covariates aside from the natural log of population represent per capita values.

Table 3.12: Event Study Coefficients

	Grant Amount	Sewerage Capital
e = -16	0.0000	-4.0571
		(1.3551)
e = -15	0.0000	0.8037
		(1.2730)

e = -14	0.0000	2.5714
		(1.4952)
e = -13	0.0000	-1.1177
		(1.2199)
e = -12	0.0000	-0.4861
		(1.2813)
e = -11	0.0000	0.9238
		(1.0574)
e = -10	0.0000	-0.1678
		(1.3288)
e = -9	0.0000	0.0719
		(1.4024)
e = -8	0.0000	-1.4882
		(1.0897)
e = -7	0.0000	1.2816
		(1.0761)
e = -6	0.0000	0.5743
		(1.7228)
e = -5	0.0000	1.4460
		(1.8526)
e = -4	0.0000	-1.7757
		(1.6682)
e = -3	0.0000	-2.1195
		(1.2274)
e = -2	0.0000	1.1961
		(0.9192)
e = -1	0.0000	0.6774
		(1.3010)
e = 0	39.8180	-0.2082
	(9.5342)	(1.3512)
e = 1	15.3706	4.3462
	(6.1208)	(2.1316)
e = 2	24.4435	10.3150
	(6.2881)	(3.5160)
e = 3	27.4015	12.5877
	(9.7786)	(3.6743)
e = 4	17.6683	12.4909
	(4.9166)	(3.9238)
e = 5	7.8658	11.9075
	(2.5698)	(3.8063)
e = 6	6.0865	10.4708
	(1.3436)	(4.0569)
e = 7	6.8891	4.5029
	(4.6979)	(3.2382)
e = 8	7.8224	3.3674
	(4.6979)	(2.8253)
e = 9	4.5610	3.1520
	(1.7787)	(2.7576)
e = 10	6.9069	3.2678
	(2.3469)	(3.0731)
e = 11	4.3978	0.3406
	(1.4327)	(2.4918)

Notes: This table presents the $\widehat{\theta}_D(e)$ from Figures 3.3 and 3.4.

Table 3.13: Pass-Through Results are Robust to Different Cutoffs

Panel A: Non-Compliant and < $138	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	7.396***	6.682	
	(.752)	(7.729)	
Pass-through			.904
			(1.047)
p-value: Pass-through = 0			0.388
p-value: Pass-through = 1			0.924
Observations	2254	2254	2254
Panel B: Non-Compliant and >= $138			
Treat * Post	47.80***	13.043	
	(6.788)	(7.337)	
Pass-through			0.273
			(0.152)
p-value: Pass-through = 0			0.073
p-value: Pass-through = 1			0.000
Observations	1426	1426	1426

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panels B and C of Table 3.4 using a different cutoff. Panel A summarizes pass-through for stacks defined by non-compliant municipalities who received grants totalling less than $138, and Panel B summarizes pass-through for all other stacks defined by non-compliant municipalities.

Table 3.14: Compliant Pass-Through is Low for Large and Small Grants

Panel A: Compliant and < $125	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	6.788***	-2.686	
	(1.825)	(1.557)	
Pass-through			-0.395
			(0.275)
p-value: Pass-through = 0			0.151
p-value: Pass-through = 1			0.000
Observations	598	598	598
Panel B: Compliant and >= $125			
Treat * Post	34.34***	6.567	
	(3.782)	(9.208)	
Pass-through			0.191
			(0.277)
p-value: Pass-through = 0			0.491
p-value: Pass-through = 1			0.004
Observations	322	322	322

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panels B and C of Table 3.4 on compliant municipalities.

Table 3.15: Pass-Through by Population Tercile

Panel A: Tercile 1	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	23.96***	8.551*	
	(8.208)	(4.596)	
Pass-through			0.357
			(0.240)
p-value: Pass-through = 0			0.137
p-value: Pass-through = 1			0.007
Observations	1584	1584	1584
Panel B: Tercile 2			
Treat * Post	18.08***	10.93***	
	(3.143)	(3.802)	
Pass-through			0.605**
			(0.279)
p-value: Pass-through = 0			0.030
p-value: Pass-through = 1			0.156
Observations	1628	1628	1628
Panel C: Tercile 3			
Treat * Post	20.58***	5.383**	
	(4.892)	(2.203)	
Pass-through			0.262
			(0.169)
p-value: Pass-through = 0			0.119
p-value: Pass-through = 1			0.000
Observations	1628	1628	1628

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results from Panel B of Table 3.2 by terciles of population. Tercile 1 has the lowest population in our sample, and Tercile 3 has the highest.

Table 3.16: Re-Weight Stacked Results to Reflect All CWA Grants

	(1) Non-Compliant	(2) Compliant	(3) Full Sample
Pass-Through	0.397	0.0337	0.353
	(0.310)	(0.363)	(0.278)
p value: Pass-through = 0	0.052	0.075	0.204
p value: Pass-through = 1	0.200	0.331	0.020
Observations	3680	920	4600

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table approximates pass-through of all CWA grants by weighting together pass-through estimates for compliant and non-compliant municipalities. Column 1 estimates $C_{it} = \alpha_0 + \beta D_{it} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its}$ on stacks defined by non-compliant municipalities. D_{it} is a dummy variable that equals one after grant receipt, we use D_{it} as a instrument for grant amount. Column 2 repeats this process for compliant municipalities. Column 3 presents a weighted average of these two coefficients where compliant municipalities represent 12 percent of the estimate and non-municipalities municipalities represent the other 88 percent.

Table 3.17: Selection on Gains

	(1) 1972-1974	(2) 1975-1977	(3) 1978-1985	(4) Full Sample	(5) Full Sample
	Sewerage Capital Spending				Grant Amount
Treat * Post	6.36	6.13	4.244	5.86***	12.62***
				(2.15)	(1.779)
Observations	3630	1656	495	3630	3630

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of $\theta_S = \sum_{g=1972}^{1985} \tilde{\theta}(g)P(G = g)$ where $\tilde{\theta}(g) = \frac{1}{1985-g+1}\sum_{t=1972}^{1985} 1\{g \leq t\}ATT(g,t)$. Column 1 presents the average of the $\tilde{\theta}(g)$ for municipalities treated in 1972-1974. Column 2 presents the average of the $\tilde{\theta}(g)$ for municipalities treated in 1975-1977, and column 3 presents the average of the $\tilde{\theta}(g)$ for municipalities treated in 1978-1981. Column 4 estimates θ_S for the entire sample and column 5 estimates the associated first stage.

Table 3.18: Full Sample Results (Adjusted for Local Matching)

	Grant Amount	Sewerage Capital	
Treat * Post	26.67***	8.84***	
	(3.783)	(2.412)	
Pass-through			0.331***
			(0.106)
p-value: Pass-through = 0			0.002
p-value: Pass-through = 1			0.000
Observations	3630	3630	3630

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panel B of Table 3.2 after adjusting grants for local matching.

Table 3.19: Grant Pass-Through for Sub-Groups (Adjusted for Local Matching))

	(1)	(2)	(3)
Panel A: Non-Compliant and < \$125	Grant Amount	Sewerage Capital	
Treat * Post	9.31***	7.23	
	(0.980)	(7.701)	
Pass-through			0.782
			(0.853)
p-value: Pass-through = 0			0.359
p-value: Pass-through = 1			0.798
Observations	2162	2162	2162
Panel B:Non-Compliant and >= \$125			
Treat * Post	61.84***	11.81	
	(8.006)	(7.343)	
Pass-through			0.191
			(0.120)
p-value: Pass-through = 0			0.112
p-value: Pass-through = 1			0.000
Observations	1518	1518	1518
Panel C: Compliant			
Treat * Post	22.20***	0.553	
	(5.619)	(4.801)	
Pass-through			0.025
			(0.286)
p-value: Pass-through = 0			0.930
p-value: Pass-through = 1			0.001
Observations	920	920	920

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Table 3.4 after adjusting grants for local matching.

Table 3.20: Average Pass-Through

	(1)	(2)	(3)
	Grant Amount	Sewerage Capital	
Treat * Post	21.73***	7.429	
	(3.487)	(5.943)	
Average Pass-through			0.212
			(1.070)
p-value: Pass-through = 0			0.843
p-value: Pass-through = 1			0.462
Observations	4600	4600	4600

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table estimates $Y_{it} = \alpha_0 + \beta D_{it} + \alpha_{is} + \alpha_{is} + \varepsilon_{its}$ on a pooled sample of all stacks. Grant amount is the dependent variable in column 1, and sewerage capital spending is the dependent variable in columns 2 and 3. Column 3 shows the average pass-through rate obtained from dividing the each stack's reduced form by its first stage.

Table 3.21: Full Sample Results (Never-Treated)

	Grant Amount	Sewerage Capital	
Treat * Post	20.58***	5.274***	
	(1.31)	(1.660)	
Pass-through			0.256***
			(0.079)
p-value: Pass-through = 0			0.001
p-value: Pass-through = 1			0.000
Observations	8492	8492	8492

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panel B of Table 3.2 using never-treated units as a control group

Table 3.22: Grant Pass-Through for Sub-Groups (Never-Treated)

	(1)	(2)	(3)
Panel A: Non-Compliant and < $125	Grant Amount	Sewerage Capital	
Treat * Post	3.53***	3.35***	
	(.230)	(1.087)	
Pass-through			0.974***
			(.310)
p-value: Pass-through = 0			0.002
p-value: Pass-through = 1			0.865
Observations	4494	4494	4494
Panel B:Non-Compliant and >= $125			
Treat * Post	22.72***	8.131***	
	(1.756)	(1.142)	
Pass-through			0.376***
			(.058)
p-value: Pass-through = 0			0.000
p-value: Pass-through = 1			0.000
Observations	3940	3940	3940
Panel C: Compliant			
Treat * Post	11.69***	1.226	
	(1.597)	(1.753)	
Pass-through			0.105
			(0.160)
p-value: Pass-through = 0			0.512
p-value: Pass-through = 1			0.000
Observations	3476	3476	3476

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Table 3.4 using never-treated units as a control group.

Table 3.23: Observable Characteristics Do Not Predict Timing of Grant Receipt

	(1)	(2)	(3)
Baseline Year:	1971	1974	1977
Compliant Facility	0.408	1.211	-0.641
	(0.777)	(1.804)	(2.137)
ln Population	-0.823**	0.0214	0.400
	(0.342)	(0.395)	(0.971)
Total Revenue	-0.00813	0.0130	0.0347
	(0.0185)	(0.0133)	(0.0356)
Total IGR	-0.00629	0.00447	0.000202
	(0.0105)	(0.0139)	(0.0264)
Federal IGR	0.00482	-0.0141	0.00558
	(0.0118)	(0.0161)	(0.0553)
Nonwater Utility Rev	0.0406	0.0116	-0.00982
	(0.0287)	(0.0249)	(0.0344)
Property Tax Revenue	0.00822	0.0297	-0.0256
	(0.00796)	(0.0227)	(0.0298)
Total Expenditure	0.00872	-0.0116	-0.0486
	(0.0177)	(0.0127)	(0.0494)
Total IG Exp	-0.0186	-0.109*	0.0598
	(0.0303)	(0.0554)	(0.0691)
Capital Outlays	0.00992	0.00255	0.0427
	(0.0145)	(0.0123)	(0.0429)
Highway Exp	-0.0141	0.00760	-0.0404
	(0.0289)	(0.0677)	(0.0723)
Sewerage Capital Outlay	-0.0120	0.0276	0.00423
	(0.0233)	(0.0239)	(0.0464)
Sewerage Non-capital Exp	0.0129	-0.0114	-0.00647
	(0.100)	(0.175)	(0.133)
Nonwater Utiltiy Exp	-0.0359	-0.00549	0.0140
	(0.0264)	(0.0208)	(0.0293)
Water IG Exp	0.159	-0.197	-0.0221
	(0.109)	(0.203)	(0.973)
Water Non-capital Exp	-0.0370	-0.0406	0.241
	(0.0758)	(0.0916)	(0.315)
Water Capital Outlay	-0.0890***	-0.0194	-0.0138
	(0.0337)	(0.0361)	(0.0897)
Water Utility Revenue	-0.0312	0.0476	-0.158
	(0.0382)	(0.0561)	(0.212)
p value	0	0	.08
Observations	112	52	35

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents results from regressing the year a municipality receives its first CWA grant on potential predictive covariates. Column 1 estimates $GrantYear_i = \alpha_0 + compliant_i + X_{ip} + \varepsilon_{ip}$ for all treated municipalities using covariate data from 1971. Columns 2 and 3 use data from 1974 and 1977, respectively, and limit the data to municipalities that do not receive a grant until after those years. All covariates aside from the facility compliance dummy and natural log of population represent per capita values. We present p values from a test of joint significance of all variables at the bottom of the table.

Table 3.24: Full Sample Results (Stacked Difference-in-Differences)

	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	21.73***	7.429	
	(3.487)	(5.943)	
Pass-through			0.342
			(0.267)
p-value: Pass-through = 0			0.200
p-value: Pass-through = 1			0.014
Observations	4600	4600	4600

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table estimates $Y_{it} = \alpha_0 + \beta D_{it} + \alpha_{is} + \alpha_{is} + \varepsilon_{its}$ on a pooled sample of all stacks. Grant amount is the dependent variable in column 1, and sewerage capital spending is the dependent variable in columns 2 and 3. Column 3 shows the implied pass-through rate from dividing the first stage in column 1 and the reduced form in column 2.

Table 3.25: Grant Pass-Through for Sub-Groups (Dynamic CS Estimates)

	(1)	(2)	(3)
Panel A: Non-Compliant and < \$125	Grant Amount	Sewerage Capital	
Treat * Post	4.820***	5.973*	
	(0.638)	(3.338)	
Pass-through			1.239
			(0.774)
p-value: Pass-through = 0			.110
p-value: Pass-through = 1			.758
Observations	1410	1410	1410
Panel B: Non-Compliant and >= \$125			
Treat * Post	53.07***	9.25***	
	(13.868)	(3.432)	
Pass-through			0.174*
			(0.092)
p-value: Pass-through = 0			.059
p-value: Pass-through = 1			0.000
Observations	1353	1353	1353
Panel C: Compliant			
Treat * Post	16.17***	7.86	
	(3.160)	(6.196)	
Pass-through			0.486
			(0.415)
p-value: Pass-through = 0			.242
p-value: Pass-through = 1			0.215
Observations	726	726	726

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panel B of Table 3.2 on sub-samples of municipalities. Panel A summarizes pass through for non-compliant municipalities that received grants totaling less than \$125, Panel B for all other non-compliant municipalities, and Panel C for all compliant municipalities.

Table 3.26: Grant Pass-Through for Sub-Groups (Simple CS Estimates)

	(1)	(2)	(3)
Panel A: Non-Compliant and < \$125	Grant Amount	Sewerage Capital	
Treat * Post	4.140***	3.139	
	(0.494)	(2.478)	
Pass-through			0.758
			(0.759)
p-value: Pass-through = 0			0.318
p-value: Pass-through = 1			0.750
Observations	1320	1320	1320
Panel B:Non-Compliant and >= \$125			
Treat * Post	52.30***	7.386**	
	(12.281)	(3.031)	
Pass-through			0.141*
			(0.077)
p-value: Pass-through = 0			0.067
p-value: Pass-through = 1			0.000
Observations	888	888	888
Panel C: Compliant			
Treat * Post	12.471***	5.728	
	(2.181)	(4.702)	
Pass-through			0.459
			(0.406)
p-value: Pass-through = 0			0.259
p-value: Pass-through = 1			0.183
Observations	693	693	693

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panel D of Table 3.2 on sub-samples of municipalities. Panel A summarizes pass through for non-compliant municipalities that received grants totaling less than \$125, Panel B for all other non-compliant municipalities, and Panel C for all compliant municipalities. We present the p value from testing for a pass-through rate of one at the bottom of the table.

Table 3.27: Pre-CWA Sewerage Capital Spending

	(1)	(2)	(3)	(4)
	Full Sample	Small Non-compliant	Large Non-compliant	Compliant
Total Sewerage Capital Spending	2,476,679	3,410,038	2,278,557	851,912
Observations	1760	752	656	352

Notes: This table shows average total sewerage capital spending from 1955 to 1971 for the full sample of municipalities, non-compliant municipalities that received grants totaling less than \$125 per capita, non-compliant municipalities that received grants totaling more than \$125 per capita, and compliant municipalities.

Table 3.28: Water Revenue Only Decreased in Compliant Municipalities

	(1) Compliant	(2) Non-Compliant
Panel A: Simple Aggregation	Water Revenue	
Treat * Post	-8.825***	0.653
	(1.131)	(0.953)
Observations	726	2904
Panel B: Dynamic Aggregation		
Treat * Post	-5.873***	0.352
	(1.161)	(0.963)
Observations	726	2904

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table presents estimates of the effect of grant receipt on water revenue. Panel A presents averages of group-time treatment effects based on group size calculated with the following equation $\frac{1}{\kappa}\sum_{g=1972}^{1985}\sum_{t=1972}^{1987} 1\{g \leq t\}\widehat{ATT(g,t)}P(G = g)$ where $\kappa = \sum_{g=1972}^{1985}\sum_{t=1972}^{1987} 1\{g \leq t\}P(G = g)$. Panel B presents averages of the $\widehat{\theta}_D(e)$ for each time relative to treatment e for $e \geq 0$ calculated with $\widehat{\theta}_D(e) = \sum_{g=1972}^{1985}\sum_{t=1957}^{1987} 1\{t-g = e\}\widehat{ATT(g,t)}P(G = g|t-g = e)$. Column 1 presents results for compliant municipalities with not-yet-treated compliant municipalities acting as controls, and column 2 presents results for non-compliant municipalities with not-yet-treated non-compliant municipalities acting as controls.

Table 3.29: Municipalities Raised Water Bills When Grants Were Too Small

	(1) Less than Costs	(2) Greater than Costs
	Water Revenue	
Treat * Post	1.1798	-0.2012
	(1.7285)	(0.8005)
Observations	960	825

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table re-estimates the results in column 2 of Table 3.28 on sub-samples of municipalities where grants were larger or smaller than estimated costs. Column 1 shows results for non-compliant municipalities where costs were greater than grant amount, and column 2 shows results for non-compliant municipalities where costs were less than grant amount.

Table 3.30: Main Water Revenue Estimates with Alternative Data

	(1) Water Revenue	(2) Water Revenue	(3) Water Revenue
Grant X Compliant	-2.500*	-2.488*	-2.316*
	(1.277)	(1.288)	(1.189)
Timing Group X Year FE	X	X	
Region X Year FE		X	
Observations	6412	6412	6412

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Table 3.9 on an alternative dataset. Column 1 presents estimates of θ from $R_{it} = \alpha_0 + \theta grant_{it} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it}$. Column 2 re-estimates the result in column 1 with region-by-year fixed effects. Column 3 re-estimates the result in column 1 with year fixed effects instead of timing-group-by-year fixed effects.

Table 3.31: Timing Water Revenue Estimates with Alternative Data

	(1) Compliant	(2) Non-Compliant
Panel A: Simple Aggregation	Water Revenue	
Treat * Post	-7.212	0.366
	(5.281)	(2.378)
Observations	1881	6363
Panel B: Dynamic Aggregation		
Treat * Post	-7.122	-1.008
	(13.493)	(2.907)
Observations	1881	6363

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table re-estimates the results in Table 3.28 on an alternative dataset.

Table 3.32: How Much Crowd-Out can Redistribution Account For?

	(1) sewerage capital	(2) water revenue	(3) sewerage capital
grant X compliant	-4.786*	-2.323**	
	(2.450)	(1.157)	
water capital			0.485
			(0.358)
grant X compliant X e	-0.512**	-0.364***	
	(0.245)	(0.138)	
water capital			0.712
			(0.433)
Observations	4752	4752	4752

Standard errors in parentheses, clustered at the municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: Panel A presents estimates of θ from $Y_{it} = \alpha_0 + \theta grant_{it} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it}$. Panel B presents estimates of θ from $Y_{it} = \alpha_0 + \theta grant_{it} * compliant_i * e_{it} + \alpha_i + \alpha_{gt} + \varepsilon_{it}$ where e_{it} indicates year relative to treatment. The dependent variable is sewerage capital spending in column 1 and water revenue in columns 2 and 3. In column 3, we use $grant_{it} * compliant_i$ as an instrument for sewerage capital spending

Table 3.33: Bacon Decomposition of Sewerage Capital Spending TWFE Estimate

	(1) Weight	(2) Average Estimate
Newly Treated vs Not-Yet-Treated	0.454	4.071
Newly Treated vs Already-Treated	0.546	2.512

Notes: This table presents the average difference-in-difference estimates and their weights from decomposing β_{rf}^{TWFE} from $C_{it} = \alpha_0 + \beta_{rf}^{TWFE} D_{it} + \alpha_i + \alpha_t + \varepsilon_{it}$ using Goodman-Bacon et al. (2019).

Table 3.34: Dose-Response TWFE Estimates

	(1) Small Non-Compliant	(2) Large Non-Compliant	(3) Compliant
	Cumulative Sewerage Capital Spending		
Cumulative Grant Dollars	1.045**	0.133	0.410*
	(0.517)	(0.100)	(0.214)
p value: Pass-through = 0	0.043	0.184	0.056
p value: Pass-through = 1	0.931	0.000	0.059
Observations	2068	1804	968

Standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of $C_{it} = C_{it} = \alpha_0 + \beta_{dr}^{TWFE} g_{it} + \alpha_i + \alpha_t + \varepsilon_{it}$ on sub-samples of municipalities. C_{it} is per capita sewerage capital expenditure, g_{it} is per capita grant amount, and α_i and α_t are municipality and year fixed effects. Column 1 summarizes pass through for non-compliant municipalities that received grants totaling less than $125, column 2 for all other non-compliant municipalities, and column 3 for all compliant municipalities.

Table 3.35: Definition of Compliance

	EPA Standard Only	Higher Standard	Full Sample
At or Above Standard	5,872	1,113	6,985
Below Standard	4,613	3,334	7,947
Total	10,485	4,447	14,932

Notes: This table

Table 3.36: Definition of Compliance for Merged Facilities

	EPA Standard Only	Higher Standard	Full Sample
At or Above Standard	1,603	579	2,182
Below Standard	2,424	1,830	4,254
Total	4,027	2,409	6,436

Notes: This table presents the results from Table 3.35 for facilities that received CWA grants.

Table 3.37: Stacked Difference-in-Difference Pass-Through Estimates (EPA Mandate Only)

	(1)	(2)	(3)
Panel A: Non-Compliant and < \$125	Grant Amount	Sewerage Capital	
Treat * Post	7.717***	9.308	
	(1.076)	(7.525)	
Pass-through			1.176
			(0.997)
p-value: Pass-through = 0			0.238
p-value: Pass-through = 1			0.865
Observations	1656	1656	1656
Panel B: Non-Compliant and >= \$125			
Treat * Post	56.71***	19.95**	
	(11.667)	(7.662)	
Pass-through			0.352**
			(0.145)
p-value: Pass-through = 0			0.015
p-value: Pass-through = 1			0.000
Observations	1150	1150	1150
Panel C: Non-Compliant and < \$167			
Treat * Post	8.808***	9.410	
	(1.215)	(7.852)	
Pass-through			1.068
			(0.971)
p-value: Pass-through = 0			0.272
p-value: Pass-through = 1			0.936
Observations	1932	1932	1932
Panel D: Non-Compliant and >= \$167			
Treat * Post	61.61***	21.18***	
	(11.851)	(7.71)	
Pass-through			0.344***
			(0.141)
p-value: Pass-through = 0			0.015
p-value: Pass-through = 1			0.000
Observations	874	874	874

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panels B and C of Table 3.4 on the sample of municipalities that did not have to satisfy state wastewater capital requirements greater than the EPA's secondary treatment requirement. Panel A summarizes pass-through for stacks defined by non-compliant municipalities who received grants totalling less than \$125, and Panel B summarizes pass-through for all other stacks defined by non-compliant municipalities. Panel C summarizes pass-through for stacks defined by non-compliant municipalities that received grants totalling less than \$167, and Panel D repeats this for all other stacks defined by non-compliant municipalities

Table 3.38: Non-Linear Least Squares Estimate of Break in Cost (EPA Mandate Only)

	(1)
	Total Cost
Split	167.3
	(237.7)
Observations	29

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table re-estiamtes the results in Table 3.3 on a sample of non-compliant municipalities that did not face state level treatment technology requirements greater than the EPA's secondary treatment standard.

Table 3.39: Full Sample Results (Nominal)

	Grant Amount	Sewerage Capital	
Treat * Post	26.31***	13.60***	
	(3.589)	(3.200)	
Pass-through			0.517
			(0.126)
p-value: Pass-through = 0			0.000
p-value: Pass-through = 1			0.000
Observations	3630	3630	3630

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panel B of Table 3.2 measuring spending and grant amount in nominal terms.

Table 3.40: Grant Pass-Through for Sub-Groups (Nominal)

	(1)	(2)	(3)
Panel A: Non-Compliant and < \$125	Grant Amount	Sewerage Capital	
Treat * Post	6.72***	9.37	
	(0.747)	(7.740)	
Pass-through			1.39
			(1.21)
p-value: Pass-through = 0			0.251
p-value: Pass-through = 1			0.745
Observations	2162	2162	2162
Panel B:Non-Compliant and >= \$125			
Treat * Post	50.75***	18.55**	
	(6.32)	(9.456)	
Pass-through			0.365*
			(0.188)
p-value: Pass-through = 0			0.052
p-value: Pass-through = 1			0.001
Observations	1518	1518	1518
Panel C: Compliant			
Treat * Post	30.39***	14.30	
	(4.622)	(8.10)	
Pass-through			0.471*
			(0.269)
p-value: Pass-through = 0			0.081
p-value: Pass-through = 1			0.049
Observations	920	920	920

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Table 3.4 measuring dollars in nominal terms.

Table 3.41: Water Revenue Results (Nominal)

	(1)	(2)	(3)	(4)
	Water Revenue	Water Revenue	Water Revenue	Water Revenue
Grant X Compliant	-5.353**			
	(2.587)			
Grant X Compliant X e		-0.850**	-0.956**	-1.233***
		(0.376)	(0.424)	(0.442)
Timing Group X Year FE	X	X	X	
Region X Year FE			X	
Observations	4752	4752	4752	4840

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Table 3.9 measuring water revenues in nominal terms.

References

Acemoglu, D., Autor, D., Dorn, D., and Hanson, G. (2014). Import competition and the great us employment sag of the 2000s. *Journal of Labor Economics*, 34.

Acton, R. K. (2021). Community college program choices in the wake of local job losses. *Journal of Labor Economics*, 39(4):1129–1154.

Ahlquist, J. and Downey, M. (2023). The effects of import competition on unionization. *American Economic Journal: Economic Policy*, 15(4).

Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, 56(4):1360–1446.

Altonji, J. G., Kahn, L. B., and Speer, J. D. (2016). Cashier or consultant? entry labor market conditions, field of study, and career success. *Journal of Labor Economics*, 34(S1):S361–S401.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Number 8769 in Economics Books. Princeton University Press.

Autor, D., Dorn, D., and Hanson, G. (2019). When work disappears: Manufacturing decline and the falling marriage market value of young men. *American Economic Review: Insights*, 1(2):161–78.

Autor, D., Dorn, D., and Hanson, G. H. (2021). On the persistence of the china shock. Working Paper 29401, National Bureau of Economic Research.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review*, 103(6):2121–68.

Autor, D. H., Dorn, D., and Hanson, G. H. (2015). Untangling trade and technology: Evidence from local labour markets. *The Economic Journal*, 125(584):621–646.

Autor, D. H., Dorn, D., Hanson, G. H., and Song, J. (2014). Trade Adjustment: Worker-Level Evidence *. *The Quarterly Journal of Economics*, 129(4):1799–1860.

Baicker, K. (2001). Government decision-making and the incidence of federal mandates. *Journal of Public Economics*, 82:147–194.

Baicker, K. and Gordon, N. (2006). The effect of state education finance reform on total local resources. *Journal of Public Economics*, 90:1519–1535.

Banzhaf, H. S. and Walsh, R. P. (2013). Segregation and tiebout sorting: The link between place-based investments and neighborhood tipping. *Journal of Urban Economics*, 74:83 – 98.

Barnes, T. (2015). 6 low-performing nashville schools to share 3$mgrant.TheTennessean$.

Barrow, L. and Rouse, C. E. (2004). Using market valuation to assess public school spending. *Journal of Public Economics*, 88(9-10):1747–1769.

Bartik, T. J. and Lachowska, M. (2014). The Short-Term Effects of the Kalamazoo Promise Scholarship on Student Outcomes. In *New Analyses of Worker Well-Being*, volume 38 of *Research in Labor Economics*, pages 37–76. Emerald Group Publishing Limited.

Baum-Snow, N. and Lutz, B. F. (2011). School desegregation, school choice, and changes in residential location patterns by race. *American Economic Review*, 101(7):3019–46.

Bayer, P., Blair, P. Q., and Whaley, K. (2020). A national study of school spending and house prices. Working Paper 28255, National Bureau of Economic Research.

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

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70(5):9–49.

Beecher, A. (2014). Tn schools extend their classroom hours. *The Tennessean*.

Beheshti, D. (2022). The impact of opioids on the labor market: Evidence from drug rescheduling. *Journal of Human Resources*.

Billings, S., Brunner, E., and Ross, S. (2017). Gentrification and failing schools: The unintended consequences of school choice under nclb. *The Review of Economics and Statistics*, 100.

Bishop, J. H. and Mane, F. (2004). The impacts of career-technical education on high school labor market success. *Economics of Education Review*, 23(4):381–402. Special Issue In Honor of Lewis C. Solman.

Black, D., McKinnish, T., and Sanders, S. (2005). The economic impact of the coal boom and bust*. *The Economic Journal*, 115(503):449–476.

Black, S. E. (1999). Do better schools matter? parental valuation of elementary education. *The Quarterly Journal of Economics*, 114(2):577–599.

Black, S. E. and Machin, S. (2011). Housing Valuations of School Performance. In Hanushek, E., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 3 of *Handbook of the Economics of Education*, chapter 10, pages 485–519. Elsevier.

Bloom, N., Kurmann, A., Handley, K., and Luck, P. (2019). The Impact of Chinese Trade on U.S. Employment: The Good, The Bad, and The Apocryphal. Technical report.

Brandt, L., Van Biesebroeck, J., Wang, L., and Zhang, Y. (2017). Wto accession and performance of chinese manufacturing firms. *American Economic Review*, 107(9):2784–2820.

Brunner, E., Hyman, J., and Ju, A. (2020). School finance reforms, teachers' unions, and the allocation of school resources. *Review of Economics and Statistics*, 102(3):473–489.

Burga, R. and Turner, S. (2022). Does enrollment lead to completion? measuring adjustments in education to local labor market shocks. *The Journal of human resources*, page 121.

Caetano, G. (2019). Neighborhood sorting and the value of public school quality. *Journal of Urban Economics*, 114(C).

Callaway, B., Goodman-Bacon, A., and Sant'Anna, P. H. C. (2021). Difference-in-differences with a continuous treatment.

Callaway, B. and Sant'Anna, P. H. (2020a). did: Difference in differences. R package version 2.0.1.906.

Callaway, B. and Sant'Anna, P. H. (2020b). Difference-in-differences with multiple time periods. *Journal of Econometrics*.

Cameron, S. and Taber, C. (2004). Estimation of educational borrowing constraints using returns to schooling. *Journal of Political Economy*, 112(1):132–182.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5):1127–1160.

Card, D., Mas, A., and Rothstein, J. (2011). *Chapter 14. Are Mixed Neighborhoods Always Unstable? Two-Sided and One-Sided Tipping*, pages 237–256. University of Pennsylvania Press, Philadelphia.

Card, D. and Payne, A. A. (2002). School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of Public Economics*, 83(1):49 – 82.

Carlson, D. and Lavertu, S. (2018). School improvement grants in ohio: Effects on student achievement and school administration. *Educational Evaluation & Policy Analysis*, 40(3):287 – 315.

Cascio, E. U. and Narayan, A. (2022). Who needs a fracking education? the educational response to low-skill-biased technological change. *ILR Review*, 75(1):56–89.

Case, K. E. and Shiller, R. J. (1989). The efficiency of the market for single-family homes. *The American Economic Review*, 79(1):125–137.

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

Cellini, S. R., Ferreira, F., and Rothstein, J. (2010). The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design*. *The Quarterly Journal of Economics*, 125(1):215–261.

Charles, K. K., Hurst, E., and Notowidigdo, M. J. (2018). Housing booms and busts, labor market opportunities, and college attendance. *American Economic Review*, 108(10):2947–94.

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

Choi, J. (2023). The effect of deindustrialization on local economies: Evidence from new england textile towns. Technical report, Princeton University.

Coelli, M. B. (2011). Parental job loss and the education enrollment of youth. *Labour Economics*, 18(1):25–35.

Collins, C. A. and Kaplan, E. K. (2017). Capitalization of school quality in housing prices: Evidence from boundary changes in shelby county, tennessee. *American Economic Review*, 107(5):628–32.

Copeland, C. (2016). Clean water act: a summary of the law. Technical report, Congressional Research Service.

Courant, P., Gramlich, E., and Rubinfeld, D. (1979). The stimulative effects of intergovernmental grants or why money sticks where it hits, in "fiscal federalism and grants-in-aid"(mieszkowski, p. and oakland, wh, eds.). *Coupe Papers on Public Economics, Urban Institute, Washington DC*.

Craig, S. and Inman, R. P. (1986). Education, welfare and the" new" federalism: State budgeting in a federalist public economy. In *Studies in state and local public finance*, pages 187–228. University of Chicago Press.

Daigger, G. T. (1998). *Upgrading wastewater treatment plants*, volume 2. CRC Press.

Deshpande, M. and Li, Y. (2019). Who is screened out? application costs and the targeting of disability programs. *American Economic Journal: Economic Policy*, 11(4):213–48.

Dhar, P. and Ross, S. L. (2012). School district quality and property values: Examining differences along school district boundaries. *Journal of Urban Economics*, 71(1):18–25.

Di Giacomo, G. and Lerch, B. (2021). Automation and human capital adjustment: The effect of robots on college enrollment. Technical report, SSRN.

Diamond, R. and McQuade, T. (2019). Who wants affordable housing in their backyard? an equilibrium analysis of low–income property development. *Journal of Political Economy*, 127(3):1063 – 1117.

Disalvo, R. (2022). Publicly-funded place-based investments and renter welfare. Technical report, Princeton University.

Dobkin, C., Finkelstein, A., Kluender, R., and Notowidigdo, M. J. (2018). The economic consequences of hospital admissions. *American Economic Review*, 108(2):308–52.

Dragoset, L., Thomas, J., Herrmann, M., Deke, J., James-Burdmy, S., Graczweki, C., Boyle, A., Upton, R., Tanenbaum, C., Jessica, G., and Wei, T. (2017). School improvement grants: Implementation and effectiveness. Technical report, U.S. Department of Education.

Eckert, F., Fort, T. C., Schott, P. K., and Yang, N. J. (2021). Imputing missing values in the us census bureau's county business patterns. Technical report, National Bureau of Economic Research.

Feenstra, R. C., Romalis, J., and Schott, P. K. (2002). U.s. imports, exports, and tariff data, 1989-2001. Working Paper 9387, National Bureau of Economic Research.

Feler, L. and Senses, M. Z. (2017). Trade shocks and the provision of local public goods. *American Economic Journal: Economic Policy*, 9(4):101–43.

Ferriere, A., Navarro, G., and Reyes-Heroles, R. (2018). Escaping the Losses from Trade: The Impact of Heterogeneity on Skill Acquisition. Technical report.

Figlio, D. N. and Lucas, M. E. (2004). What's in a grade? school report cards and the housing market. *American Economic Review*, 94(3):591–604.

Fisher, R. C. and Papke, L. E. (2000). Local government responses to education grants. *National Tax Journal*, 53(1):153–168.

Flynn, P. and Marcus, M. M. (2021). A watershed moment: The clean water act and infant health. Working Paper 29152, National Bureau of Economic Research.

Flynn, P. and Smith, T. (2021). When and why do intergovernmental grants crowd-out local spending? evidence from the clean water act. Technical report, Vanderbilt University.

Flynn, P. and Smith, T. (2022). Rivers, lakes and revenue streams: The heterogeneous effects of clean water act grants on local spending. *Journal of Public Economics*, 212:104711.

Foote, A. and Grosz, M. (2020). The Effect of Local Labor Market Downturns on Postsecondary Enrollment and Program Choice. *Education Finance and Policy*, 15(4):593–622.

Friday, C. (2021). The impact of drastic school interventions on academic performance. Technical report, Vanderbilt University.

Friday, C. and Smith, T. (2023). Turning around schools (and neighborhoods?): School improvement grants and gentrification. *Economics of Education Review*, 94:102382.

Gamkhar, S. (2000). Is the response of state and local highway spending symmetric to increases and decreases in federal highway grants? *Public Finance Review*, 28(1):3–25.

Gamkhar, S. (2003). The role of federal budget and trust fund institutions in measuring the effect of federal highway grants on state and local government highway expenditure. *Public Budgeting & Finance*, 23(1):1–21.

Gamkhar, S. and Oates, W. (1996). Asymmetries in the response to increases and decreases in intergovernmental grants: Some empirical findings. *National tax journal*, pages 501–512.

Gibbons, S., Machin, S., and Silva, O. (2013). Valuing school quality using boundary discontinuities. *Journal of Urban Economics*, 75(C):15–28.

Ginsburg, A. and Smith, M. S. (2018). Revisiting sig: Why critics were wrong to write off the federal school improvement grant program. Technical report, FutureEd.

Gonzales, J. (2015a). 6 metro nashville public schools get improvement grants. *The Tennessean*.

Gonzales, J. (2015b). Metro changes principals at 5 struggling schools. *The Tennessean*.

Gonzales, J. (2015c). Turnaround principal announced for napier elementary. *The Tennessean*.

Goodman, J., Hurwitz, M., and Smith, J. (2017). Access to 4-year public colleges and degree completion. *Journal of Labor Economics*, 35(3):829–867.

Goodman-Bacon, A. (2021a). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225.

Goodman-Bacon, A. (2021b). Drdecomp: Stata module to perform a bacon decomposition of dose-response difference-in-differences estimation.

Goodman-Bacon, A. (2021c). Online appendix: Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277. Themed Issue: Treatment Effect 1.

Goodman-Bacon, A., Goldring, T., and Nichols, A. (2019). Bacondecomp: Stata module to perform a bacon decomposition of difference-in-differences estimation.

Gordon, N. (2004). Do federal grants boost school spending? evidence from title i. *Journal of Public Economics*, 88(9-10):1771–1792.

Gramlich, E. M. (1969). State and local governments and their budget constraint. *International Economic Review*, 10(2):163–182.

Greenland, A. and Lopresti, J. (2016). Import exposure and human capital adjustment: Evidence from the U.S. *Journal of International Economics*, 100(C):50–60.

Greenland, A., Lopresti, J., and McHenry, P. (2019). Import Competition and Internal Migration. *The Review of Economics and Statistics*, 101(1):44–59.

Gross, T., Notowidigdo, M. J., and Wang, J. (2020). The marginal propensity to consume over the business cycle. *American Economic Journal: Macroeconomics*, 12(2):351–84.

Grossman, J. B. and Roberts, J. (1989). Welfare savings from employment and training programs for welfare recipients. *The Review of Economics and Statistics*, pages 532–537.

Grosz, M. (2022). Do postsecondary training programs respond to changes in the labor market? *Journal of Human Capital*, 16(4):461–487.

Grosz, M., Kurlaender, M., and Stevens, A. (2022). Capacity and flexibility in community college cte programs: Program offerings and student success. *Research in Higher Education*, 63.

Guerrieri, V., Hartley, D., and Hurst, E. (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics*, 100:45 – 60.

Hecker, D. (2005). High-technology employment: A naics-based update. *Monthly labor review / U.S. Department of Labor, Bureau of Labor Statistics*, 128:57–72.

Heckman, J. J. and Mosso, S. (2014). The economics of human development and social mobility. *Annual Review of Economics*, 6(1):689–733.

Helm, I. and Stuhler, J. (2021). The dynamic response of municipal budgets to revenue shocks.

Henderson, J. M. (1968). Local government expenditures: A social welfare analysis. *The Review of Economics of Statistics*, pages 156–163.

Hershbein, B. and Stuart, B. (2023). The evolution of local labor markets after recessions. Technical report, Federal Reserve Bank of Philadelphia.

Hershbein, B. J. (2012). Graduating high school in a recession: Work, education, and home production. *The B.E. Journal of Economic Analysis Policy*, 12(1).

Hilger, N. G. (2016). Parental job loss and children's long-term outcomes: Evidence from 7 million fathers' layoffs. *American Economic Journal: Applied Economics*, 8(3):247–83.

Hines, J. R. and Thaler, R. H. (1995). The flypaper effect. *Journal of economic perspectives*, 9(4):217–226.

Hoxby, C. M. and Kuziemko, I. (2004). Robin hood and his not-so-merry plan: Capitalization and the self-destruction of texas' school finance equalization plan. Working Paper 10722, National Bureau of Economic Research.

Hoynes, H., Miller, D. L., and Schaller, J. (2012). Who Suffers during Recessions? *Journal of Economic Perspectives*, 26(3):27–48.

Huttunen, K. and Riukula, K. (2019). Parental Job Loss and Children's Careers. IZA Discussion Papers 12788, Institute of Labor Economics (IZA).

Hyman, B. G. (2022). Can Displaced Labor Be Retrained? Evidence from Quasi-Random Assignment to Trade Adjustment Assistance. Working Papers 22-05, Center for Economic Studies, U.S. Census Bureau.

Inman, R. P. (2008). The flypaper effect. Working Paper 14579, National Bureau of Economic Research.

Jackson, C. K. (2010). A little now for a lot later. *Journal of Human Resources*, 45(3):591–639.

Jackson, C. K. (2018). Does school spending matter? the new literature on an old question. Working Paper 25368, National Bureau of Economic Research.

Jerch, R. L. (2018). The local consequences of federal mandates: Evidence from the clean water act. Technical report, Working paper, Johns Hopkins University.

Johnson, R. C. and Jackson, C. K. (2019). Reducing inequality through dynamic complementarity: Evidence from head start and public school spending. *American Economic Journal: Economic Policy*, 11(4):310–49.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics*, 17(2):303–316.

Kane, T. J., Riegg, S. K., and Staiger, D. O. (2006). School quality, neighborhoods, and housing prices. *American Law and Economics Review*, 8(2):183–212.

Keiser, D. A., Kling, C. L., and Shapiro, J. S. (2019). The low but uncertain measured benefits of us water quality policy. *Proceedings of the National Academy of Sciences*, 116(12):5262–5269.

Keiser, D. A. and Shapiro, J. S. (2019). Consequences of the clean water act and the demand for water quality. *Quarterly Journal of Economics*.

Kho, A., Henry, G. T., Pham, L. D., and Zimmer, R. (2020). Spillover effects of recruiting teachers for school turnaround: Evidence from tennessee. *Educational Evaluation and Policy Analysis*, 0(0):01623737221111807.

Knight, B. (2002). Endogenous federal grants and crowd-out of state government spending: Theory and evidence from the federal highway aid program. *American Economic Review*, 92(1):71–92.

Kovalenko, A. (2023). Natural resource booms, human capital, and earnings: Evidence from linked education and employment records. *American Economic Journal: Applied Economics*, 15(2):184–217.

Kuka, E., Shenhav, N., and Shih, K. (2020). Do human capital decisions respond to the returns to education? evidence from daca. *American Economic Journal: Economic Policy*, 12(1):293–324.

Laajaj, R., Moya, A., and Sánchez, F. (2022). Equality of opportunity and human capital accumulation: Motivational effect of a nationwide scholarship in colombia. *Journal of Development Economics*, 154:102754.

Lafortune, J., Rothstein, J., and Schanzenbach, D. W. (2018). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics*, 10(2):1–26.

Leduc, S. and Wilson, D. (2017). Are state governments roadblocks to federal stimulus? evidence on the flypaper effect of highway grants in the 2009 recovery act. *American Economic Journal: Economic Policy*, 9(2):253–92.

Lee, M. J. (2021). The effect of import competition on educational attainment at the postsecondary level: Evidence from nafta. *Economics of Education Review*, 82:102117.

Liu, K., Salvanes, K. G., and Sørensen, E. (2016). Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession. *European Economic Review*, 84:3–17. European Labor Market Issues.

Lochner, L. and Monge-Naranjo, A. (2012). Credit constraints in education. *Annual Review of Economics*, 4(1):225–256.

Londoño-Vélez, J., Rodríguez, C., and Sánchez, F. (2020). Upstream and downstream impacts of college merit-based financial aid for low-income students: Ser pilo paga in colombia. *American Economic Journal: Economic Policy*, 12(2):193–227.

Lutz, B. (2010). Taxation with representation: Intergovernmental grants in a plebiscite democracy. *The Review of Economics and Statistics*, 92(2):316–332.

Lutz, B. (2015). Quasi-experimental evidence on the connection between property taxes and residential capital investment. *American Economic Journal: Economic Policy*, 7(1):300–330.

Machin, S. (2011). Houses and schools: Valuation of school quality through the housing market. *Labour Economics*, 18(6):723–729.

McGrew, W. (2019). U.s. school segregation in the 21st century. *Washington Center for Equitable Growth*.

McGuire, M. (1978). A method for estimating the effect of a subsidy on the receiver's resource constraint: with an application to us local governments 1964–1971. *Journal of Public Economics*, 10(1):25–44.

McKinnish, T., Walsh, R., and White, T. K. (2010). Who gentrifies low-income neighborhoods?. *Journal of Urban Economics*, 67(2):180 – 193.

Megdal, S. B. (1987). The flypaper effect revisited: An econometric explanation. *The Review of Economics and Statistics*, pages 347–351.

Mello, S. (2019). More cops, less crime. *Journal of Public Economics*, 172:174–200.

Mian, A., Rao, K., and Sufi, A. (2013). Household Balance Sheets, Consumption, and the Economic Slump*. *The Quarterly Journal of Economics*, 128(4):1687–1726.

Miller, S., Wherry, L. R., and Foster, D. G. (2023). The economic consequences of being denied an abortion. *American Economic Journal: Economic Policy*, 15(1):394–437.

Mountjoy, J. (2022). Community colleges and upward mobility. *American Economic Review*, 112(8):2580–2630.

National Environmental Research Center (1972). Upgrading existing wastewater treatment plants. Technical report.

Oates, W. (1969). The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the tiebout hypothesis. *Journal of Political Economy*, 77(6):957–71.

Oates, W. E. (1999). An essay on fiscal federalism. *Journal of economic literature*, 37(3):1120–1149.

Oreopoulos, P., von Wachter, T., and Heisz, A. (2012). The short- and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics*, 4(1):1–29.

Pham, L. D., Henry, G. T., Kho, A., and Zimmer, R. (2020). Sustainability and maturation of school turnaround: A multiyear evaluation of tennessee's achievement school district and local innovation zones. *AERA Open*, 6(2):2332858420922841.

Pierce, J. R. and Schott, P. K. (2016). The surprisingly swift decline of us manufacturing employment. *American Economic Review*, 106(7):1632–62.

Pierce, J. R. and Schott, P. K. (2020). Trade liberalization and mortality: Evidence from us counties. *American Economic Review: Insights*, 2(1):47–64.

Pierce, J. R., Schott, P. K., and Tello-Trillo, C. (2022). Trade Liberalization and Labor-Market Outcomes: Evidence from US Matched Employer-Employee Data. Working Papers 22-42, Center for Economic Studies, U.S. Census Bureau.

Rambachan, A. and Roth, J. (2023). A More Credible Approach to Parallel Trends. *The Review of Economic Studies*. rdad018.

Roth, J. and Sant'Anna, P. H. (2021). Efficient estimation for staggered rollout designs. *arXiv preprint arXiv:2102.01291*.

Rowley, S. (1993). China woos western businesses, snubs clinton.

Rubin, K. (1985). Efficient investments in wastewater treatment plants. US Congress, Congressional Budget Office.

Ruhm, C. J. (1991). Are Workers Permanently Scarred by Job Displacements? *American Economic Review*, 81(1):319–324.

Salvanes, K. G., Willage, B., and Willén, A. L. (2021). The Effect of Labor Market Shocks across the Life Cycle. Technical report.

Schueler, B. E., Asher, C. A., Larned, K. E., Mehrotra, S., and Pollard, C. (2020). Improving low-performing schools: A meta-analysis of impact evaluation studies. Technical Report 274, Annenberg Institute at Brown University.

Schwandt, H. and von Wachter, T. (2019). Unlucky cohorts: Estimating the long-term effects of entering the labor market in a recession in large cross-sectional data sets. *Journal of Labor Economics*, 37(S1):S161–S198.

Schwartz, A. E., Voicu, I., and Horn, K. M. (2014). Do choice schools break the link between public schools and property values? Evidence from house prices in New York City. *Regional Science and Urban Economics*, 49(C):1–10.

Smith, J., Goodman, J., and Hurwitz, M. (2020). The economic impact of access to public four-year colleges. Working Paper 27177, National Bureau of Economic Research.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374.

Strumpf, K. S. (1998). A predictive index for the flypaper effect. *Journal of Public Economics*, 69(3):389–412.

Stuart, B. A. (2022). The long-run effects of recessions on education and income. *American Economic Journal: Applied Economics*, 14(1):42–74.

Sun, M., Kennedy, A., and Loeb, S. (2020). The longitudinal effects of school improvement grants. (177).

Sun, M., Penner, E. K., and Loeb, S. (2017). Resource- and approach-driven multidimensional change: Three-year effects of school improvement grants. *American Educational Research Journal*, 54(4):607 – 643.

Superfine, B. M. (2011). Stimulating school reform: The american recovery and reinvestment act and the shifting federal role in education. *Mo. L. Rev.*, 76:81.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64(5):416–424.

Travis, R., Morris, J. C., and Morris, E. D. (2004). State implementation of federal environmental policy: Explaining leveraging in the clean water state revolving fund. *Policy Studies Journal*, 32(3):461–480.

Urban-Brookings Tax Policy Center (2017). State and local finance data: Exploring the census of governments. Technical report.

US Census Bureau (2021). About annual survey of state and local government finances. Technical report.

USDA (2019). High energy cost grants fact sheet. Technical report.

USDA (2020). Solid waste management grants fact sheet. Technical report.

USEPA (1973). Estimate of municipal wastewater treatment facility requirements. Technical report, Clean Watershed Needs Survey.

USEPA (1980). Handbook of procedures: Construction grants program for municipal wastewater treatment works. Technical report.

USEPA (1998). How wastewater treatment works. *Technical Report*.

USEPA (2000). Progress in water quality: an evaluation of the national investment in municipal wastewater treatment. *Technical Report*.

USEPA (2002). The clean water and drinking water infrastructure gap analysis. *Technical Report*.

USGAO (1994). Water pollution: Information on the use of alternative wastewater treatment systems. Technical report.

Vegh, C. and Vuletin, G. (2015). Unsticking the flypaper effect in an uncertain world. *Journal of Public Economics*, 131.

Weinstein, R. (2022). Local labor markets and human capital investments. *Journal of Human Resources*, 57(5):1498–1525.

Wigger, C. (2020). Decoupling homes and schools assessing the impact of forced school choice on residential change. Technical report, Northwestern University.

Yinger, J., Bloom, H. S., and Boersch-Supan, A. (2016). *Property taxes and house values: The theory and estimation of intrajurisdictional property tax capitalization*. Elsevier.

Zheng, A. (2019). Residential sorting, school choice, and inequality.

Zimmer, R., Henry, G. T., and Kho, A. (2017). The effects of school turnaround in tennessee's achievement school district and innovation zones. *Educational Evaluation & Policy Analysis*, 39(4):670 – 696.