ESSAYS IN LABOR ECONOMICS

By

Priyankar Datta

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics – Doctor of Philosophy

2022

ABSTRACT

ESSAYS IN LABOR ECONOMICS

By

Priyankar Datta

This dissertation contains three chapters that study the impact of a labor market policy on nursing home staffing and patient outcomes, the impact of parental divorce on long-term market outcomes, and the impact of a change in housing wealth on children's schooling decisions.

Chapter one examines the effect of paid sick leave mandates on nursing home outcomes, with a focus on low-paid nursing staff. I use the synthetic control group method and traditional difference-in-differences models along with Nursing Home Compare data and Vital Statistics microdata to estimate the causal effect of paid sick leave mandates on nursing home outcomes. I find significant increases in part-time nursing assistant staffing and resident health and safety improvements. Nursing homes in areas with sick pay mandates also show reductions in the elderly mortality rate. Nursing assistant hours per resident day increase by 2.3 percent driven by a 12 percent increase in the hours for part-time workers, and there are no significant reductions in hours of full-time nursing assistants. I find improvements along multiple measures of patient health and safety. My calculations show that sick pay mandates helped prevent at least 4000 nursing home deaths per year among the elderly.

Chapter two explores the importance of divorce in explaining the gender gap in children's long-term educational outcomes. I find large differences in the gender gap between divorced and non-divorced families. Boys perform much worse in divorced families. I use a sibling fixed effects model to find that boys in divorced families have a lower likelihood of graduating high school and attending college relative to their sisters. My results show that boys' likelihood of graduating high school declines by 6.4 percentage points if their parents are divorced before they turn 13, and their chances of attending college decline by 12.2 percentage points if they

are a teenager at the time of divorce. I find that parents' divorce is unrelated to the gender gap in achievement scores. My event study models show a drop in boys' achievement scores relative to girls around the time of divorce.

Chapter three examines the effect of housing wealth changes on private school enrolment. I use data from The National Longitudinal Survey of Youth's child supplement to examine the relationship between housing wealth and private school enrolment. I use a multinomial logit model and find that self-reported housing price changes increase the likelihood that respondents switch from private to public school. Heterogeneity analyses reveal that house price increases have a positive relationship between switching from private to public school across income, gender, race, and religion. Finally, a rise in house prices increases the likelihood that a child moves from public school to private school when transitioning from middle school to private school.

ACKNOWLEDGEMENTS

I want to thank my parents, Chaitali and Prasenjit Datta, for the significant sacrifices they made in raising me and for their continuous support during my education. I would also like to thank my dissertation committee members, Scott Imberman, Todd Elder, and Ben Zou, who were generous in giving me their time for extensive feedback and thoughtful guidance. Finally, I want to thank Michigan State University for creating an ideal community that always strives for higher targets.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

THE IMPACT OF MANDATED PAID SICK LEAVE LAWS ON THE LONG-TERM CARE INDUSTRY

1.1 Introduction

In 2019, less than a third of workers in the bottom 10 percent of the income distribution had access to paid sick leave (PSL), compared to 90 percent of workers in the top quarter of the income distribution (BLS 2020). Nursing assistant is one of the lowest paid occupations in the United States and in 2019 had a median wage of about 14 dollars, with almost 34 percent exposed to increases in the minimum wage from 2014-2018 (Ruffini 2020). Information regarding paid sick leave availability among nursing assistants is varied[1]. However, most reports point to the fact that there are many nursing assistant jobs that do not have paid sick leave (Dill et al. 2013). Data from the National Health Interview Survey covering 2014-2018 show that only 55 percent of nursing assistants receive paid sick leave. As higher-wage earners are more likely to be covered by an employer's PSL policy when not mandated by law, universal requirements can help level the playing field.

The United States and Japan are the only two industrialized nations without universal access to PSL. The recent Family First Coronavirus Response Act[2], provides up to two weeks of paid sick time at 100 percent of the person's salary; however, the legislation is temporary and expired at the end of 2020. Sick workers are less productive than those at full health (Goetzel et al. 2004) moreover, coming to work sick can be especially risky in a fragile setting like the long-term care industry. This makes nursing assistants and their patients especially vulnerable. When facilities do not offer separate paid time off for sickness, those with access

[1]Analysis based on data from a 2004 survey Squillace et al. (2009) points to 70 percent of nursing assistant having access to paid leave. However, the survey does not separate between paid time off and paid sick leave.

[2]This act exempts health care workers, however, the HEROES Act, passed on May 15, removes this exemption.

to paid vacation or unallocated paid time off may be reluctant to use it when sick, instead preferring to save it for other uses. Most nursing assistants are women, minorities, have less than a college education, and come from low-income households. When faced with whether to stay at home or come to work when sick, a difficult financial position may increase their likelihood of choosing the unhealthy and unproductive option of coming in to work.

This paper explores the effect of state-mandated PSL laws on the long-term care industry, with a particular focus on nursing assistants and residents. I examine how these laws affect nursing assistant staffing, patient conditions, and elderly nursing home mortality rate. I calculate the age-adjusted elderly mortality rate from Vital Statistics microdata and use objective measures of patient health and safety from the near universe of nursing homes in the United States. I estimate the effect of a PSL mandate on these outcomes by exploiting the temporal and spatial variation in the implementation of city and statewide PSL mandates. Some nursing assistants may already have PSL, which may improve nursing quality, but state-mandated PSL may lead nursing homes to increase the amount of sick time they offer to their staff. Conversely, PSL mandates may lead nursing homes to reduce staff or the salaries paid, which may reduce the quality of care. Nursing homes may also want to keep staff hours constant by increasing the number of part-time staff available to substitute for possible increases in leave taking by full-time staff. Since the effects are theoretically ambiguous, the answer to how this policy will affect nursing home quality must be determined empirically.

I use two well-established empirical methods in the applied microeconomics literature to examine the effect of mandated PSL laws on the quality of care in nursing homes. First, I use traditional difference-in-differences models to estimate the effect of PSL mandates on nursing assistant staffing and nursing home patient conditions. Second, I use the synthetic control group method (SCGM) to estimate the effect of PSL mandates on nursing home elderly mortality rate. The setting of this paper is well-suited for the application of the SCGM. First, I can match the elderly nursing home mortality rate of the treated units for a long pre-reform period. Second, the regions are heterogeneous in terms of size, nursing home mortality rate, and population density and thus provide broad common support. As a result, the findings should have external validity for other U.S. areas with a similar nursing home structure and policy environment. The rich temporal and spatial variation in PSL mandates allows for comparing patient well-being due to changes in non-wage labor costs, along with flexibly accounting for demographic and economic changes at a small geographic level. I leverage nursing home level information from the Centers for Medicare and Medicaid's (CMS) OSCAR (Online Survey Certification & Reporting System) and CASPER (Certification And Survey Provider Enhanced Reports) data sets along with death counts from all county multiple cause of death microdata to look at a variety of measures widely considered to be indicators of facility quality and patient well-being.

I show that PSL mandates do not lead to a decrease in full-time staffing hours but an increase in part-time nursing assistant hours. This is consistent with existing literature on the effects of PSL mandates on aggregated labor markets (Pichler and Ziebarth (2019); Stearns and White (2018)). Increased part-time staff hours lead to an increase in average nursing assistant hours per resident. I find that nursing assistants' time spent with patients increased by 2.3 percent following the enactment of PSL mandates. Nursing homes in areas with PSL mandates show a 12 percent increase in part-time nursing assistant hours compared to those in areas without mandated PSL laws.

If PSL mandates reduce the likelihood of presenteeism [3], health care quality should improve. In order to explore this further, I look at the impact of PSL mandates on common health care quality measures for nursing homes. I find that PSL mandates improve patient health and reduce the severity of safety violations. Specifically, I show that PSL mandates reduce the share of residents with pressure ulcers and anti-psychotic medications by 12 percent and 5 percent, respectively. I also find that severe violations are reduced by 7.5 percent. I perform event study analyses and do not find strong evidence of preexisting trends in any outcome. Further, after the mandate, I observe persistent improvements in

[3] Presenteeism refers to the lost productivity that occurs when employees are not fully functioning in the workplace because of an illness, injury, or other condition.

patient health and nursing home staffing. I also provide evidence of how PSL mandates affect the elderly nursing home mortality rate using the SCGM. My analysis shows the pre-reform outcome dynamics of the treated group to closely match that of the synthetic control group, thus providing support for valid counterfactuals. I find that PSL mandates lead to a 5.5 percent fall in the elderly nursing home mortality rate, translating to a reduction in 4479 deaths.

The nursing home industry serves almost 1.6 million elderly residents across the United States. Despite the CMS introducing several regulations to ensure the quality of care, there are still significant variations in the quality of care provided. It may be the case that lower-quality nursing homes are the ones where the PSL mandates are binding and have the most potent effects. Exploring possibilities of heterogeneous outcomes, I find that most improvements in care quality are driven by outcomes from nursing homes with a high share of Medicaid patients. Enactment of PSL mandates also leads to nursing homes changing patient composition along several margins, including healthier patients and a slight increase in private-paying patients. However, I find no significant reduction in the share of Medicaid or Medicare paying patients.

Contagious presenteeism[4] behavior by aides and orderlies can be life-threatening for patients but can be potentially minimized by paid sick time. This study's findings provide evidence of how PSL mandates can impact the long-term care industry. More than half of individuals reaching the age of 65 will require long-term care at some point in their lives, much of which is provided in residential settings (Favreault and Dey 2016). The long-term care industry accounts for a substantial portion of the U.S. economy, making up more than 10 percent of the total expenditure on Medicare and Medicaid. As the elderly share of the U.S. population continues to grow, the effect of paid sick leave mandates on nursing home patient well-being is an increasingly important policy question.

[4]Contagious presenteeism is when employees with a contagious disease (e.g., a common cold) go to work sick and spread the disease to co-workers, customers, and the general population.

1.1.1 Research on Sick Leave

State-mandated PSL is a recent phenomenon in the United States. Literature on the effects of such mandates is new and growing. There does not exist any work examining the effect of such mandates on the nursing home industry. Nonetheless, multiple studies are looking at how PSL mandates impact workers' productivity, employment, and health outcomes. This paper is the first to examine PSL mandates' effects on the nursing home industry. It is an industry where a significant portion of workers are low paid, do not have access to PSL, and attending work ill can have considerable adverse effects.

This paper contributes to multiple strands of literature in economics and public health. Existing literature finds mixed evidence of the effects of the mandates on employment. Connecticut was the first state to implement the policy in 2012, and reports have found reductions in annual hours worked (Ahn and Yelowitz 2016) along with increased unemployment and economically insignificant changes in labor force participation (Ahn and Yelowitz 2015). However, more recent work looking at a longer timeline and multiple states has found no significant reductions in employment or wage growth (Pichler and Ziebarth 2019).

This paper also contributes to a small but growing literature studying the effects of PSL mandates on worker productivity. A survey of New York City businesses found that the large majority of businesses observed no effect on productivity, with only 2 percent reporting that productivity had increased and 4 percent reporting that productivity decreased (Appelbaum and Milkman 2016) from PSL mandates. Additionally, a survey in Jersey City found that more than a third of businesses noticed improvements in productivity (Lindemann and Britton 2015). Worker satisfaction has also improved following the enactments of PSL mandates. More than half of employees in San Francisco who previously had access to PSL reported improved employer support and an increase in the number of sick days provided (Drago et al 2011). Productivity of workers also improves through improved health outcomes of workers. The lowest-income group of workers without paid sick time were at the highest risk of delaying and forgoing medical care for themselves and their family members (DeRigne et al.

2016). PSL mandates have also been shown to reduce aggregate illness-related leave taking (Stearns and White 2018). Thus, existing literature points to substantial public health externalities of PSL mandates through the reduced spread of illness and disease to coworkers and customers.

1.1.2 Research on the Long-Term care Industry

This paper contributes to the public health literature focusing on the long-term care industry. Increased staffing through laws and business cycle changes have been shown to reduce violations and mortality (Chen and Grabowski (2015); Matsudaira (2014); Park and Stearns (2009); Antwi and Bowblis (2018); Stevens et al. (2015)). Reduced staffing through unionization does not harm patient outcomes, which shows that labor policies may influence worker productivity (Sojourner et al., 2015). I find that PSL mandates increased time spent by staff on patients, similar to results achieved from wage increases and increased staff attention (Grabowski et al. 2011). Changes in staff turnover driven by macroeconomic fluctuations have also been shown to reduce mortality and the number of violations (Antwi and Bowblis (2018); Stevens et al. (2015). Other state and city level policies like minimum wages have also been shown to improve (Ruffini 2020) patient health and safety outcomes in nursing homes.

The remainder of this paper proceeds as follows. Section 2 describes the nursing home industry and paid sick leave laws. Section 3 outlines the data. Section 4 describes the empirical framework. Section 5 presents results, and Section 6 describes the robustness checks, and Section 7 concludes.

1.2 Institutional Details

1.2.1 Nursing Homes

The United States has almost 16,000 nursing homes that provide round-the-clock care to their residents. The 1.4 million residents of nursing homes receive health, personal care, and supportive and rehabilitative services. The vast majority of these residents are 65 years or older, with a significant number being 80 years or older. They receive routine assistance in daily activities ranging from eating, bathing, dressing, mobility, and toileting (Centers for Medicaid and Medicare Services 2015). Due to the relatively inelastic demand for their service, most nursing homes have very high occupancy rates. These facilities are also highly labor intensive and employ nearly 1.6 million workers, with around 40 percent being nursing assistants (Ruffini 2020). Due to the considerable time nursing assistants spend with elderly and fragile residents, their tasks can directly affect patient well-being. They provide basic patient care under the direction of nursing staff. Their primary duties are feeding, bathing, dressing, grooming, moving patients, and changing linens (ONET 2019).

Nursing homes have to fulfill several federal reporting and inspection requirements. The 1987 Nursing Home Reform Act (NHRA), requires annual independent health inspections; nursing credentialing, minimum RN staffing levels, and routine, comprehensive patient assessments (Castle and Ferguson (2010); Institute of Medicine (1986)). Fulfilling these requirements makes nursing homes eligible to receive Medicare and Medicaid reimbursement. The nursing home market has significant barriers to entry: certificate of need laws places limits on construction and the number of beds each facility can have in many states (NCSL (2019); Centers for Medicaid and Medicare Services (2015)). Only 26 percent of residents pay out of pocket for nursing home stays, with the rest coming from Medicare (12 percent) and Medicaid(62 percent) reimbursement. Medicaid reimbursement rates are 30 percent lower than Medicare's on average and are roughly half of the out-of-pocket prices. These reimbursement rates are set by expected patient costs, with Medicaid rates depending on state payment structures and Medicare rates on service needs and local cost-of-living adjustments (Houser et al. (2018); Centers for Medicaid and Medicare Services (2019)). Residents paying out of pocket are generally much more responsive to quality and prices than those covered by public insurance (Gertler 1989).

1.2.2 Paid Sick Leave

The Family and Medical Leave Act of 1993 (FMLA) is the only existing federal law that provides sick leave. However, it is relatively restrictive compared to recent local mandates providing only unpaid leave to employees with at least 1250 hours worked annually at a business with greater than or equal to 50 employees (Tominey 2016). The restrictive nature of this bill leaves out 49 million workers, almost 44 percent of all employees (Jorgensen and Appelbaum 2014). Table A.1 provides a summary of the mandates effective on or before 2018. The details of the bills differ by jurisdiction, but nearly all sick pay mandates are employer mandates.

The first PSL mandate requiring employers to provide paid sick days was implemented in San Francisco in 2007. Connecticut became the first state in the U.S. to enact PSL legislation in 2012. The Connecticut law mandated that firms with 50 or more employees offer paid sick time to service workers; the San Francisco policy included no such exemptions based on firm size or industry. I study PSL mandates from 9 states and 10 localities for nursing home staffing and nursing home level patient conditions. I study PSL mandates from Connecticut, Massachusetts, California, and New York City to study the effect of these mandates on elderly nursing home mortality. These states span several regions, from New England (Connecticut, Massachusetts, Vermont, Rhode Island), the West (California, Oregon, Washington, Arizona), and the Mid-Atlantic (Maryland, New Jersey)[5]. The localities with PSL mandates in my sample consist primarily of large cities and counties located in the above states, as well as New York City; Philadelphia, PA; Minneapolis and St. Paul, MN; Chicago and Cook County, IL. Twenty-two states have passed preemption laws preventing localities from requiring employers to provide PSL. These include four states that concurrently passed statewide laws prohibiting localities from establishing PSL requirements that differ from existing state standards.

Almost all states mandate a sick leave accrual rate between 1 and 1.3 hours per 40 hours.

[5]The most recent state to enact a PSL mandate is Michigan, coming in to effect March 2019

Some localities cap the amount of sick time that can be accrued, often tying the limit to the employer's size. A preponderance of localities specify that paid sick time can be used for reasons related to domestic violence or sexual assault and to care for oneself or a family member. Virtually all local laws include exemptions, many of which relate to the number of hours an employee works. For instance, the law in Cook County, IL, exempts employees who work less than 80 hours a year. There are also laws exempting healthcare workers; for instance, Washington DC's 2009 law exempts healthcare workers and is not part of my sample. Long-term care workers in Vermont and New Jersey who work per diem are also exempt from the law.

1.3 Data

My primary data sources are the online survey certification and reporting (OSCAR) and certification and survey provider enhanced reporting (CASPER) data from the CMS. Nursing facilities must report staffing numbers and patient characteristics to CMS to be eligible for Medicaid and Medicare reimbursement. During the analysis period, these data are based on staffing numbers for the two weeks before unannounced health and safety inspection (Centers for Medicare and Medicaid 2020). For 2000-2018, the OSCAR/CASPER data provide two measures of employment for nursing assistants: hours per resident per day and the number of full-time equivalent (FTE) staff by part-time, full-time, and contractor status. Assuming full time staff work 35 hours a week, the number of FTE full-time staff is the total hours worked by full-time staff in the week, divided by 35. Full-time staff are defined as those working at least 35 hours a week, and part-time are those working fewer than 35 hours a week. Staffing hours per resident day are provided by the Brown School of Public Health (2020), denoted as the number of FTE multiplied by 35 from the OSCAR/CASPER data, divided by the number of residents in the facility and then processed to account for implausibly large year-to-year fluctuations in staffing levels.

As with staffing information, facilities are required to report information on patient conditions to CMS to be eligible for reimbursement. These assessments are conducted by facility staff and are subject to a CMS audit. My analyses focus on the fraction of residents with conditions most likely to be affected by the quality and quantity of nursing care: moderate-to-severe pressure ulcers; urinary tract infections (UTI); physical restraints; or psychotropic medication. I focus only on long-term stays (residents in a facility for at least 100 days), as these patients have the most prolonged exposure to a facility's nursing staff.

State surveyors conduct unannounced health inspections every 9-15 months on nursing homes and interview staff, patients, and family members about the quality of care (Associates Inc and Abt Associates Inc. (2013); Centers for Medicaid and Medicare Services (2015)). OSCAR/CASPER data has the type, number, severity, and scope of each violation a facility has received and the date the inspection occurred. Several violations are closely associated with patient safety measures and measures of worker productivity like routinely assessing residents, communicating patient conditions to family members, changing bed linens, avoiding accident hazards, and providing sanitary food preparation.

Following Ruffini (2020), I use every violation a facility has received since 2000 and construct several patient safety measures. I consider the total number of violations and the number of severe violations that present immediate harm or danger to residents. I create two measures of violations following Ruffini (2020). The first considers all health violations. The second considers Quality of Care(QOC) violations. The QOC measure includes violations in the assessment relating to the quality of care, nursing, dietary, physician, rehabilitative services, dental, and pharmacy regulation categories. These violations are the subset of violations widely recognized in the public health literature to be most closely related to nursing responsibilities (Chen and Chen (2019), Harrington et al. (2000), Harrington et al. (2002), Antwi and Bowblis (2018)). Violations are not uncommon in nursing homes, and almost all nursing homes have at least one recorded violation every year, with the average nursing home having 7 violations. Depending on the scope and severity, violations can lead to substantial fines and penalties. Reduced violations may lead to significant cost savings for nursing homes.

Average patient age, the share of female residents, and other demographic variables used are available from the Minimum Data Set and provided through Brown School of Public Health (2020). The main specifications also control for county-level demographic and economic controls that change over time. Total and elderly population figures are available through the National Institute of Health (2020). To account for local labor market conditions, I control for the overall county unemployment rate using data from Bureau of Labor Statistics (2020). Finally, to ensure my results are not driven by the overall state policy environment or other policy changes coincident with PSL mandates, I control for state EITC parameters, the share of the elderly population receiving Supplemental Security Income (SSI), a proxy for Medicaid eligibility, and AFDC/TANF caseloads and benefit levels from the University of Kentucky (2020). The share of Medicaid claimants at the establishment level and private ownership and chain status is provided through Brown School of Public Health (2020). My sample has slightly more than 15,000 facilities and around 3000 counties over 19 years.

Finally, I calculate state-year and county-year mortality rates for elderly nursing home residents using death counts from Vital Statistics micro-record multiple cause of death files, and county-by-age population counts from the National Institute of Health (2020). I follow the methodology first discussed in Stevens et al. (2015) that adjusts the outcome to create a measure of mortality rate that holds the age distribution constant over time. Nursing home deaths are identified as those occurring in nursing home/long-term care centers. The age-adjusted mortality rate measure at the county-level[6] is defined as follows:

$$m_{cy} = \Sigma_{a=65}^{85+} \frac{deaths_{cay}}{pop_{cay}} * \frac{pop_{a,2010}}{\Sigma_{k=65}^{85+} pop_{k,2010}} \tag{1.1}$$

[6]I calculate the age-adjusted state-level mortality rate by aggregating the population and death counts from the county level to the state level

Where $deaths_{cay}$ is the number of deaths in nursing home settings among individuals aged a in county c in year y from the Vital Statistics data, and pop_{cay} is the population size of individuals aged a in each county-year from the National Institute of Health (2020). The fraction containing the two population estimates is the national fraction of individuals age a in the elderly population in year 2010. I only consider areas with at least 3 post-treatment periods for the mortality analysis. Detailed information on areas that mandated paid sick leave laws on or before 2015 is provided in Table A4. I also restrict my main analysis to the four treatment regions of California, Connecticut, Massachusetts, and New York City. All 4 of these areas have much higher levels of nursing home beds than the remaining smaller areas. For example, Connecticut has the smallest number of beds among my main treatment areas, but it still has almost four times the number of beds as the next largest area Philadelphia. This distinction is important as small changes in the number of deaths in an already narrowly defined outcome variable can cause large relative changes in the outcome variable. In Table A.19, I provide results for additional counties.

Table A.2 reports information on the control variables used in my analysis. I find large differences along several policy variables like minimum wage and state EITC. Facilities in treatment areas are also larger on average and face lower levels of competition. My treatment counties are, on average, more urban and have a younger population. These localities, predominantly located along the coasts, also largely vote democratic and have seen many changes in labor laws and safety nets, which explains the higher minimum wages and level of state EITC benefits.

1.4 Empirical Framework

1.4.1 Difference-in-Differences Approach

I estimate the effect of paid sick leave mandates on nursing home staffing and patient outcomes using a difference-in-differences identification strategy that exploits the temporal and geographic variation in the enactment of PSL mandates. The relationship between PSL mandates and my outcomes of interest are formalized as follows:

$$Y_{ft} = \alpha_{+}\gamma PSL_{ft} + \theta Z_{ct} + \beta X_{ft} + \delta_f + \mu_t + \epsilon_{ft} \tag{1.2}$$

PSL is an indicator for the enactment of a PSL mandate in year t[7], X is a vector of facility-level characteristics (average resident age, for profit or non-profit, chain or single establishment, total number of beds, percentage of female residents, percentage of Medicaid payors), Z is a vector of time-varying county-level factors (unemployment rate, elderly share of the population, share of state SSI recipients who are elderly, AFDC and TANF caseloads, minimum wages, state EITC rate, and degree of competition among nursing home), δ_f is a facility fixed effect, and μ_t is a year fixed effect. In my sample, PSL mandates are enacted at the city, state, and county levels. Hence, I define my variable at the establishment level. This represents a standard difference-in-differences analysis where outcomes in my treatment regions (i.e., counties and states enacting a PSL mandate) are compared to control regions with no PSL laws. I cluster my standard errors at the county level in all analyses. My dependent variable, Y_{ft} in Equation (2) represents one of several possible outcomes in nursing homes.

I find that PSL laws increase the likelihood of firms exiting the market only slightly. This suggests to the extent that my findings are not driven by low-performing firms exiting the market. The primary assumption under which my identification rests is that the trends in outcomes among facilities not receiving the treatment accurately measure counterfactual trends among the treated facilities. Figures 1, 2, 3, and 4 present event study analyses from the three main groups of outcome variables. These figures provide visual evidence of an absence of pre-trends. This lends support to my difference-in-differences identification strategy.

The traditional differences-in-difference research design with two way fixed effects (TWFE) and multiple time periods has several limitations that have been pointed out in recent papers (Borusyak and Jaravel (2017), Abraham and Sun (2018), Goodman-Bacon (2019), de Chaisemartin (2018)). In a TWFE regression, units whose treatment status does not change

[7] I define enactment of a PSL mandate as the year in which the mandate took effect

over time serve as the comparison group for units whose treatment status does change over time. This is problematic when it adjusts the path of outcomes for newly treated units by the path of outcomes for already treated units. However, this is not the path of untreated potential outcomes; it includes treatment effect dynamics. Thus, these dynamics appear in the coefficient, making it very hard to give a clear causal interpretation.

Moreover, this issue can have potentially severe consequences. For example, it is possible to come up with examples where participating in the treatment is positive for all units in all periods, but the TWFE estimation procedure leads to estimating a negative effect of participating in the treatment. Even when negative weights can be ruled out, and the coefficient recovers a weighted average of ATTs, these weights are hard to interpret.

1.4.2 The Synthetic Control Group Method

To assess the causal effect of sick pay mandates on nursing home elderly mortality rate, I use the SCGM developed by Abadie and Gardeazabal (2003). The all-county multiple causes of death Vital statistics microdata does not identify the specific nursing facility where the deceased passed away. Hence, my analysis for this outcome variable happens at an aggregated geographic level of the county or state. This creates an ideal setup for an analysis using the SCGM. The SCGM creates a weighted average of multiple control units to create a synthetic control group whose pre-reform outcome matches that of the treatment group (Abadie et al. 2010). Given certain assumptions, the differences in postreform outcomes between the treatment and the synthetic control group yield evidence of causal reform effects (Athey and Imbens 2017).

In this study, following Table A.17, the treatment units are counties or states that implemented sick pay mandates before 2015; the potential control units consist of the remaining U.S. counties or states. Because I analyze each treatment unit separately, the notation below refers to a single treatment and J control units.

Let y_{it}^0 denote the outcome that would have been observed in region i at time t in the absence of the sick pay mandate. Moreover, y_{it}^1 denotes the outcome for the treated region i at time t, where the sick pay mandate was implemented at time $T_0 + 1$. I assume $y_{it}^1 = y_{it}^0 \forall t = 1, ..., T0, \forall i = 1, ..., J + 1$.

According to (Abadie et al., 2010) the counterfactual y_{it}^0 is represented by the following factor model:

$$y_{it}^0 = \delta_t + \theta_t Z_i + \lambda_t \mu_i + \epsilon_{it} \tag{1.3}$$

where δ_t is a common time effect, θ_t is a vector of possibly time-dependent coefficients, λ_t is a vector of unobserved common factors, and μ_i is a vector of unknown factor loadings.

The first assumption relevant to synthetic control methods requires that the treatment does not affect mortality rates in control regions. This implies the absence of spatial spillover effects. In my analysis, two of three treated states have donor states contributing to the synthetic control group that are neighboring states. Table A5 lists the donor states used to generate the synthetic control unit for each treated state. The table shows that the state of Nevada and the state of New Hampshire both contribute less than 10 percent (1.8 percent and 8.6 percent, respectively) to the treated states of California and Massachusetts.

Second, similar to traditional difference-in-differences models, no unobserved shocks should affect the outcome differently for treatment and control groups in post-reform periods. In this case, shocks violating this assumption would be other policies that are correlated with sick pay mandates in treated regions (but not in control regions). The SCGM may be more suitable than traditional methods in incorporating these shocks as the synthetic control units are, by construction, built to mimic the outcome of the treated unit (which includes unobservables affecting the outcome).

1.4.2.1 Implementation

SCGM requires the estimation of two matrices: 1) V is the weighting matrix determining the relative predictive power of Z_i and of y_{it}^0 , and 2) W is a vector of nonnegative weights attached to the J control countries.

I follow the literature and calculate the weighting matrix by minimizing the mean squared prediction error (MSPE) for prereform periods:

$$MSPE = \frac{\Sigma_t (y_t^1 - y_t^0 W^*(V))^2}{T_0} \tag{1.4}$$

where T_0 represents the number of prereform time periods.

1.4.2.2 Treatment Effects and Inference

Following (Abadie et al. 2010) I calculate both the pre and post-reform MSPE and the ratio between the two. The ratio between the post and pre-reform MSPE indicates the size of a possible treatment effect. As the MSPE ratio measures the relative treatment effect, I also calculate a percentage and level treatment effect.

To conduct inference, I run placebo estimates as suggested by Abadie et al. (2010). I construct placebo estimates for each treated unit I have considered in my SCGM analysis. I follow Abadie et al. (2010), and use the rank of the true treatment estimate relative to the N placebo estimates to determine the p-value of the H_0 hypothesis of no treatment effect ($H_0 : MSPERatio_{Treat} \leq MSPERatio_{Placebo}$). This can also be thought of as a percentile rank $p = \hat{F}(MSPERatio_e)$ for the event e, where $\hat{F}$ stands for the empirical cumulative distribution of all MSPE Ratios, from the placebo estimates. I use all the individual p-values and follow the procedure used by Dube and Zipperer (2015) to calculate joint p-values based on the sum of the single p-values using the Irwin-Hall distribution.

1.5 Results

1.5.1 Staffing

Recent literature has focused on the impact of PSL mandates in the United States on employment; however, to my knowledge, there are no papers focusing on this particular market.

Using staffing data from the CMS nursing home compare and ltc focus from the Brown School of Public Health (2020), I calculate the effect of PSL mandates on nursing home staffing. I present results for nursing assistant hours per resident day and the total weekly hours by employee type per resident.

Figure 1 shows event study diagrams using the TWFE analysis on hours per resident day, the number of part-time nursing assistants per resident, and the number of contract nursing assistants per resident. The graphs in panels (a) - (c) show a sharp on impact increase in nursing assistant staffing for all three measures.

Table A.3 reports a statistically significant increase in nursing assistant hours per resident day, driven by increases in part-time nursing assistant hours. I find a very slight and statistically insignificant decrease in weekly hours of full-time nursing assistants per resident. I estimate that mandated PSL leads to a 0.356 unit increase in weekly part-time nursing assistant hours per resident in nursing homes and a 0.035 unit increase in weekly hours for contract workers. Compared to baseline estimates, this corresponds to around a 12 percent increase in weekly hours of part-time workers per resident and an almost 20 percent increase for contract workers. This drives the 2.3 percent increase in nursing assistant hours per resident day. Increased availability of sick time may lead to nursing assistants taking more time off, which may reduce attendance. Thus, nursing homes may want a cushion for potential nursing assistant leave taking by increasing part-time nursing assistant hiring. As part-time nursing assistants accrue leave at a lower rate and generally have lower privileges, this may be a way nursing homes try to counter potential staffing decreases due to leave-taking. Employers may also be looking to reduce over-time hours by hiring part-time staff, thus trying to offset possible increases in cost caused by the PSL mandates. No significant changes in the full-time nursing assistant hours may be due to reduced sickness-related leave taking by other staff not being infected due to presenteeism behavior of sick workers. The increased presence of part-time nursing staff and no changes in other staff lead to significant improvements in the net hours per resident per day. Higher hours per resident day clearly show improved quality at nursing homes.

1.5.2 Patient Safety

As PSL mandates lead to increased time spent with patients and higher part-time and contract nursing assistant hours, they may likely improve patient outcomes.

Panel A of Table A.4 shows that PSL mandates significantly reduce the number of severe violations. The results from the table translate to a 7 percent reduction in the number of severe violations. Panel B reports an 8.5 percent reduction in the number of severe quality of care violations. Similarly, existing literature on the effect of other policies on patient health shows that higher minimum wages and increases in Medicare reimbursement rates in the early 2000s reduced the number of violations and the likelihood of violations (Ruffini, 2020). My results are also similar to other studies finding changes in staffing through turnover (Antwi and Bowblis, 2018) and local unemployment (Huang et al., 2019), resulting in a significant reduction in the number of deficiencies in a nursing home.

1.5.3 Patient Health

Figure 2 shows event study diagrams using the TWFE analysis on the share of residents with pressure ulcers, the share of residents on physical restraints, and the share of residents being administered antipsychotic medications. The graphs in panels (a) and (c) show a persistent fall in the share of residents having pressure ulcers and the share of residents being administered antipsychotic medications. These show different trends in outcomes compared to the staffing event studies that show an on impact change in staffing levels. Panel (b) shows a sharp on impact increase in the share of residents with physical restraints.

Tables 3 and 4 point to improved staffing outcomes and a reduction in the number of severe deficiencies. Improvements in quality along these measures are usually expected to be associated with improved patient health measures. To explore this argument, I analyze several measures of patient health that facilities submit to CMS every quarter. Previous literature points to the fact that these measures have a strong relationship to service quality provided by direct care staff (Brandeis et al. (1994); Dorr et al. (2005); Cawley et al. (2006); Grabowski et al. (2011)). More than 5 percent of residents in my sample have pressure ulcers, a preventable health condition. As nursing assistants help residents with day-to-day activities and monitor their health, close attention can reduce the likelihood of developing ulcers. Column (1) of Table A.5 reports that PSL mandates reduce the share of residents with pressure ulcers by 0.6 percentage points, which translates to an almost 12 percent reduction in the share of pressure ulcers compared to baseline estimates.

The modal cause of bacterial-related hospitalization among nursing home residents is UTIs. Nursing assistants administer and monitor indwelling catheters, which frequently cause UTIs. Timely removal and minimizing the usage of these devices can reduce the likelihood of infection (Saint (2000), CDC (2009)). From columns (3) and (4), I find that PSL mandates reduce the likelihood of UTIs by a negligible amount.

Existing literature does not point to a clear direction in which nursing homes change the usage of physical restraints in response to higher staffing costs. However, following the results from previous tables, it may be the case that nursing homes are also providing better care along this dimension. Conversely, nursing homes may also be looking to offset increased costs from PSL by increasing the use of physical restraints. As physical restraints reduce movement, greater nursing assistant presence should reduce their usage (Cawley et al., 2006). Increased staff presence or assembly can also increase the use of these devices (Grabowski et al., 2011). Column (5) looks at the relationship between PSL mandates and physical restraint usage. I find that PSL mandates increase the usage of physical restraints by 0.8 units, translating to a 35 percent increase in usage compared to baseline estimates.

Antipsychotic medications are often used as a chemical restraint on residents with behavioral problems. They were introduced as a quality measure beginning in 2011 and have a smaller analysis sample. These drugs are primarily sedatives that may strongly affect a patient's mental processes. In 2008 the FDA issued a warning that all antipsychotics are associated with an increased risk of death for persons with dementia. In 2012 the CMS started a three-pronged strategy to reduce the unnecessary use of antipsychotics by initiating a national partnership with providers, launching an educational website, and public reporting of data on antipsychotic use. Existing literature has found that higher licensed nurse staffing leads to reduced usage of antipsychotic medication (Grabowski et al. 2011). Column (7) reports that PSL mandates reduce the usage of antipsychotic medication by 5 percent compared to baseline estimates.

Nursing homes often manage residents' behavioral health problems using labor, medications, and physical restraints. It is widely believed that antipsychotic medications are a substitute for nursing care, while physical restraint usually requires more attention. My findings of higher staffing numbers along with lower usage of psychotropic drugs and a higher share of patients with restraints are in line with Grabowski et al. (2011), where they show a 10 percent increase in wages leads to an increase in psychotropic usage between 1.1 percent to 3.5 percent and a decrease in physical restraints between 26 percent and 28 percent.

1.5.4 Patient Composition

If nursing homes change the composition of residents, it may have significant social welfare implications. Nursing homes may decide to offset increased costs borne from PSL mandates by increasing the share of patients from relatively wealthy private-paying sources. Medicaid reimbursement rates are lower than private-paying rates, which are lower than Medicare rates. In Tables (6) and (7), I look at the effect of PSL mandates on the characteristics of admitted residents. Columns (1) and (2) of panel A in Table A.6 show that PSL mandates do not lead nursing homes to significantly change the share of residents who pay via Medicare or Medicaid. I find that PSL mandates increase the share of private paying residents by 5.3 percent. I find increases in the percentage of residents with higher average care needs, shown here by the assisted daily living index and average care index, which go up 2 percent and 1.7 percent, respectively. Higher labor costs may lead to changes in facilities' discharge or admission practices. Facilities may cycle through patients by sending them to the hospital and then readmitting them to the facility, thereby gathering higher Medicare reimbursement fees. Column (4) of Table A.6 shows that facilities do not respond to PSL mandates by increasing resident churn. I do not see significant increases in occupancy rates after the enactment of PSL mandates.

Table A.7 looks at the effects on patient demographics and care needs, with Columns (1) and (2) showing characteristics that admission decisions cannot influence. I also find a significant reduction in the share of female residents by almost 1 percent. I find that nursing homes have a slightly healthier composition of patients, with fewer hospital admits by 1.1 percent. Columns (3)-(6) look at care needs that can be altered by assessor judgment. I find favorable changes in the care mix for residents, with the share with bladder incontinence decreasing significantly by 3.5 percent.

1.5.5 Mortality

Higher mortality can be an extreme negative consequence of poor health conditions in nursing homes. Nursing home residents have a much higher mortality rate than the general population of the same age. One-third of nursing home residents die within the first year of nursing home admittance, about three times the rate of the general population over 85.

The results from the difference-in-differences regression analysis on mortality rate at the county level are presented in Table A.8. The result shows a 3.2 % fall in nursing home elderly mortality rate in counties that enacted the PSL mandates.

The results from my SCGM analysis on age-adjusted elderly nursing home mortality rate are shown in Figures 5 and 6 and Table A.9. The left column of Figure 5 shows the evolution of state-level elderly nursing home mortality rates in three treatment states and county-level elderly nursing home mortality rates with the five boroughs of New York City aggregated as one treatment county. In the left column of Figure 5, the solid lines represent the treatment areas, and the dashed lines represent the synthetic control areas. The composition of each treatment state- the weights W of the J control states- are in Table A5. The solid vertical line on the x-axis represents the year when the sick pay mandate went into effect.

The right column of Figure 5 displays the permutation inference using placebo tests graphically. Following the convention in the literature, the graphs plot the differences in the mortality rate (solid black) along with the differences for all placebo SCGM runs (gray). As seen, for pre-reform periods, the solid black line fluctuates very closely around the horizontal zero line implying that the synthetic control units very closely map the mortality rate of the treatment units. After the reform, indicated by the black dashed vertical line, mortality rate differentials between treated and control areas separate, and the treated areas show a persistent negative difference with the synthetic group. One exception is New York City, where I witness an initial drop followed by a slight rise above the synthetic group. New York City is different from the other treatment areas in my analysis, the treatment area is entirely urban, orders of magnitude denser in population, and the mortality rate is at a much lower baseline level.

Figure 5 and Table A.9 show substantial differences in mortality rates between treatment areas. Column (1) of Table A.9 shows New York City has an average pre-reform mortality rate of 0.007, whereas Massachusetts has a pre-reform mortality rate of 0.016. The pretreatment elderly nursing home mortality rate of the treatment areas closely matches those of the synthetic control areas, supporting that SCGM produces valid counterfactuals. Visually, I also find sizeable and systematic reform-related drops in the elderly nursing home mortality rate in the treated states. The mortality rate of treatment states appears to differ substantially from the mortality rate of synthetic control states. The pretreatment outcome dynamics of New York City have a better match with its synthetic control group than the treated states. The SCGM analysis of New York City is at the county level and uses a donor pool of 100 counties resulting in a wider set of potential outcomes to create a weighted synthetic group.

The left column of Figure 6 shows the difference between the treated and the synthetic control group's postreform mortality rate to be substantial for all the treated states. The right column of Figure 6 provides a visual exhibition of the procedure I use to conduct inference, following Abadie et al. (2010). After plotting the ratio of post-reform MSPE to pre-reform MSPE of the treated and donor areas in descending order from the top, I find all treated states to be among the top 5 units, with Massachusetts having a rank of 1.

To quantitatively evaluate the SCGM fit between treated and controls and to conduct inference, I show all relevant results in Table A.9. Column (2) of Table A.9 shows the post-treatment MSPE to the pre-treatment MSPE. The MSPE ratios lie between 10 for California and 61 for Massachusetts. I conduct inference using the method described in Section 4.2.2. For the states, I use all never treated states between 2000 and 2018 as my donor states. Additionally, I drop the state of New York from the state-level analysis as New York City makes up 43 percent of the state's population. For New York City I choose 100 counties that provide the best pre-reform fit from all the never treated counties. To conduct inference, I replicate the standard SCGM procedure with each placebo state or county, pretending it had been treated at the same time as the actual treatment county or state. Column (5) illustrates the calculation of the p-values for the hypothesis $H0$: $MSPERatio_{Treat} \leq MSPERatio_{Placebo}$, which is simply the rank of the MSPE Ratio of the treated county or state divided by #Total Counties/States Assessed. As seen in column (4), except for California (p=0.125) all treatment areas are statistically significant, with Massachusetts having a p-value of 0.025.

I also calculate the sum of all p-values and then evaluate their joint p-value based on the Irwin-Hall distribution (Dube and Zipperer, 2015). I find a highly precise joint p-value for the effect of paid sick leave mandates on the elderly nursing home mortality rate. Column (4) of Table A.9 shows the percentage treatment effect for each treatment unit and the average for all treated units. As seen in the figures, all 3 treated states have strong negative treatment effects, with Massachusetts having a treatment effect of 8 percent. My results show a combined average negative treatment effect of 5.5 percent on elderly nursing home mortality[8]. My estimates are slightly larger than other labor market changes affecting nursing home mortality. For example, a 10 percent increase in the minimum wage reduces nursing home mortality by 3.1 percent (Ruffini, 2020) and a 1 percentage point increase in the unemployment rate decreases nursing home mortality by 4.7 percent (Stevens et al., 2015). Column (5) of Table A.9 shows the level treatment effect of elderly nursing home mortality rate for the postintervention periods. Column (6) of Table A.9 lists the percentage difference in mortality rate between the treatment and the synthetic group in 2016. Column (6) values are then used in Column (7) to calculate the net change in elderly nursing home deaths for each treatment area. I find that paid sick leave laws in the selected areas prevented an estimated 4479 deaths in total, with the bulk of reductions coming from Massachusetts and California.

1.5.6 Heterogeneity

The PSL mandate's average effects may hide the presence of heterogeneous effects. In tables 10, 11, and 12, I look at the effect on health outcomes, violations, and staffing by splitting the sample by profit status, multi-facility, and the share of residents paying via Medicaid. Table A.12 looks at the effect of PSL mandates on resident health measures by facility type. I do not find any heterogeneous effects on the share of residents with pressure ulcers and physical restraint use by facility type. For-profit and high-Medicaid share nursing homes show large drops in the share of residents on antipsychotic medications.

Panel A of Table A.11 does show that the reductions in severe violations are driven by facilities with a higher share of patients paying via Medicaid. Nursing homes with a high share of Medicaid paying residents show a more than 11 percent drop in the number of severe violations, almost ten times that of those with a low share of Medicaid payors. Panel B reports that chain nursing homes drive the decrease in the number of severe care violations. Multi-establishment nursing homes see a more than 15 percent reduction in the number

[8]Results displayed in Table A.19 show the average percentage treatment effect increasing only slightly to 5.8 percent after adding King County, WA and San Francisco County, CA

of severe quality of care violations. Table A.10 shows that increases in nursing assistant hours per resident day are significant in for-profit, multi-establishment, and high-Medicaid facilities. The estimates are much stronger for these types of facilities. Additionally, high-Medicaid facilities show a 2.6 percent increase in the hours per resident day.

Residents paying through Medicaid provide lower revenues and smaller margins. This may make them more likely to be present in lower-quality nursing homes. My data does not allow me to identify whether the high-Medicaid share nursing homes provided paid sick leave to their employees or not. However, if that is the case, PSL mandates may benefit the low-income group of an already vulnerable population. Existing literature finds for-profit nursing homes driving improved deficiency scores and lower deficiencies in times of higher unemployment rate (Huang et al. 2019). For-profit and higher-Medicaid share nursing homes were also more likely to have deficiencies in the existing literature (Harrington et al. 2000).

1.6 Robustness Checks

A potential confounding factor is that higher labor costs may cause low-performing firms to exit the market. If low-performing establishments were to exit the market, this would attenuate the aggregate benefits of PSL mandates. I construct a balanced panel of facilities that appear at any point in my sample and use my main regression equation with an outcome variable being an indicator equal to 1 if that facility appears in that given year. Table A.13 shows a slight (1.9 percent) reduction in the likelihood that a facility exists after the introduction of PSL mandates. The precision and size of the estimates lend support to my estimates not being driven by the exit of low-performing firms.

The primary assumption for my identification is that the trends in outcomes among areas not receiving the treatment accurately measure counterfactual trends among the treated areas. I provide evidence that this assumption holds in the data via event study models. I run event study analyses that are depicted in Figures (1) through (4). I use the timing of the enactment of the PSL mandates to carry out an event study analysis of the primary outcome variables. This regression model also allows me to assess whether PSL mandates lead to an on-impact change in outcomes. The event studies are generalized difference-in-differences models similar to Jacobson et al. (1993). Instead of my regular difference-in-differences specification, I use dummy variables for each period before and after enactment of the laws for a maximum of 4 periods on either side (leaving out the period before enactment). I do not find any strong indications for the presence of preexisting trends for my assessed outcome variables. This lends support to my difference-in-differences identification strategy. Figures 1 and 2 show that the effects of the mandates on health and staffing outcomes persist long after enactment.

I estimate the effect of PSL mandates on many nursing home outcome variables. It is plausible that some of these variables may be correlated with one another. This creates the issue of multiple hypothesis testing. I provide evidence through Bon-ferroni and Sidak p-values that my outcomes are mostly robust to multi-collinear problems. The p-values are provided in Tables A14 to A16. The health outcomes and staffing variables are highly significant, while the number of severe violations is just insignificant at the 10 percent level. These p-values are conservative measures, and I expect my estimates to be precise to a large extent.

1.7 Conclusion

This paper finds that state-mandated PSL leads to changes in nursing home staffing and patient well-being. The estimates are precise and meaningful. Nursing homes may be increasing the number of part-time nursing assistants to tackle possible increases in leave taking by full-time staff. However, I do not witness any significant decreases in full-time nursing assistant hours, which, combined with increased part-time hours, lead to a net increase in staff hours per resident. Nursing homes also show improvements in patient health and safety, particularly reductions in the share of patients with pressure ulcers and severe violations. Nursing homes also substitute for more attentive and less potentially harmful means of care for elderly patients. High-Medicaid share nursing homes primarily drive these results. Additionally, I find sharp and persistent decreases in nursing home elderly mortality rates after the enactment of PSL mandates. My calculations show PSL mandates prevented about 4000 elderly nursing home deaths.

Nursing assistants may also be more attentive and productive at work if PSL reduces their likelihood of showing up sick. Cost savings from improved patient outcomes can offset a significant amount of the increased cost due to PSL. As these mandates positively impact vulnerable sections of society, they have considerable potential social welfare implications. Nursing assistants and Medicaid insurance holders are more likely to come from low-income backgrounds and thus may receive greater welfare weights from the social planner. As the population of the United States ages, these policies will have significant ramifications.

CHAPTER 2

PARENTAL DIVORCE AND THE GENDER GAP IN CHILDREN'S LONG TERM LABOR MARKET OUTCOMES

2.1 Introduction

The gender gap in educational attainment has reversed substantially over the past four decades. This reversal in the gender gap has happened in many wealthy countries, including the United States. On average, 84 percent of young (25-34) women have attained at least an upper secondary level of education compared to 81 percent of young men (OECD 2013). Family disruption leads to the lower educational attainment of children. Family disadvantage has been shown to have a causal effect on the gender gap in test scores and behavioral outcomes (Autor et al. 2019). These emerging gender gaps suggest a reason for concern. Education has become an increasingly important determinant of lifetime income. The employment prospects of low-educated youths have declined sharply in recent years. Little is known about how divorce affects the gender gap in long-term educational outcomes. In this paper, I focus on divorce as a potential explanation for the gender gap in long-term labor market outcomes. My analysis looks at how long-term outcomes of children are impacted by divorce before the age of 13. Since children in single-parent families are more likely to live with their mothers, the lack of a male role model in single-parent families can adversely affect boys relative to girls (Cobb-Clarke and Tekin 2011). Divorce may differentially affect the long-term labor market outcomes of boys and girls for several reasons. First, these outcomes are more elastic to family circumstances for boys than girls (Bertrand and Pan 2013). Second, differential parental investment in girls relative to boys varies positively with age in single-parent families (Bibler 2019).

As a motivating example for my analysis, consider Figure 1, which plots the median math (Peabody Individual Achievement Mathematics test) and reading (Peabody Individual

Achievement Reading and Recognition test) test scores for children from divorced families at different periods before and after divorce. Boys start with a higher median mathematics test score than girls. However, as they approach the event of divorce, the gap gets narrower and eventually reverses in favor of girls. The eventual outcome of the gap in reading scores is similar, but the patterns from the figure reveal a different path. While boys do start with a higher median reading score compared to girls, that advantage quickly disappears as they move closer to the event of divorce. This is due to a sharp fall in the average reading score for boys and a consistent rise in girls' reading scores. Past work has shown that, on average, girls do better than boys in terms of reading as they age, which is the opposite for math scores (Barnard-Brak et al. 2015). The figure shows that boys reading scores fall while their math scores do not rise sufficiently at the median as they age. These trends motivate the hypothesis that divorce is more consequential for the educational outcomes of boys. The effects of a divorce may start affecting children's outcomes years before the event. This may be why boys' reading scores start dropping around the time of their parent's divorce. Other factors like behavioral and non-cognitive outcomes may also have a role to play in a child's long-term outcomes.

This paper seeks to quantify the contribution of divorce to the gender gap in long-term educational outcomes. To carry out this analysis, I tackle two important obstacles, lack of suitable data and credible identification. I address the data challenge using the child supplement of the national longitudinal survey of youth (CNLSY). The CNLSY tracks all children born to the women of the NLSY-1979 from 1984. The children born post-1984 are tracked from birth to adulthood. The NLSY is a nationally representative sample of youth who were 16-22 in 1979. These longitudinal data sets offer remarkable detail on family characteristics and early educational outcomes, including achievement tests for math and english, high school graduation, college graduation, college attendance, criminal convictions, self-reported health, and labor market outcomes.

The second obstacle to my study is that divorce is intrinsically confounded with outcomes independent of their impact on family environment. For example, highly educated parents are disproportionately likely to have stable marriages, enroll their children in higher quality schools, and may have children with above-average latent ability. My challenge is to separate the direct impact of divorce from the hereditary confounds that would lead to biased outcomes among children without any causal effect of divorce on children's outcomes. My empirical approach removes these confounds by contrasting the outcomes of opposite-sex siblings linked to the same mother. This provides valid identification of the differential effect of divorce on boys relative to girls. It is based on the assumption that any difference in educational outcomes that girls may have before divorce relative to their male siblings is not systematically larger or smaller relative to non-divorced families. I offer confirmations of the plausibility of this assumption through placebo tests and event study analyses. The event study analyses are helpful for two reasons. First, they show the effect of divorce on achievement scores.

Second, they allow me to check if there are any existing pre-trends. I make use of scores in various subjects as well as a measure of the home environment. In all these cases, the outcomes differ systematically by gender and marital status. However, in no case is the brother-sister gap in these outcomes predicted by whether or not parents eventually get divorced. Brothers and sisters appear equally advantaged or disadvantaged by families' future marital status.

My analysis shows that the timing and intensity of the impact of divorce differ across children. Looking at divorce through a cutoff point of age 13 may identify the effect of being divorced at a young age versus an older age. To address this challenge, I run two kinds of exposure time models. First, I use exposure time as a linear model. This captures the impact of being exposed to divorce for a year before turning 18. Second, I use age categorizations to examine divorce's impact on different age groups. Some respondents are not interviewed in all waves of the data. Including them in my sample may bias my estimates. To overcome this, I am running two different sets of regressions. First, I drop all individuals who have ever been incarcerated. Second, I include individuals who appear in at least eighty percent of the waves. The NLSY has also changed young adult interview rules over different waves. This increased attrition in certain waves. In both cases, I find no significant difference in outcomes caused by attrition.

The rest of the paper is organized as follows: Section 2 reviews the literature. Section 3 describes the Data. Section 4 discusses the empirical strategy for the primary cutoff model. Section 4 discusses the results for the primary cutoff model. Section 5 discusses the placebo tests. Section 6 presents the exposure time models. Section 7 presents the event study models. Section 8 concludes.

2.2 Literature Review

This paper contributes to several strands of literature. The literature studying the gender gap in educational outcomes is vast and active. Buchmann and DiPrete (2006), Autor et al. (2019), and Autor and Wasserman (2013) look at how gender gaps in educational outcomes are affected by different family structures. Autor et al. (2019) assess whether family disadvantage exerts a differential effect on the educational outcomes of boys relative to girls. The study finds that boys in disadvantaged families have more disciplinary problems, lower achievement scores, and are less likely to graduate high school. They provide evidence that a sizeable portion of the documented minority-white difference in educational and behavioral gender gaps is attributable to higher degrees of family disadvantage among minority families. My paper contributes to this literature by examining divorce's impact on long-term educational outcomes.

Prior research contrasts boys and girls born to different mothers to assess the relationship between family disadvantage and the gender gaps in child outcomes. The estimates generated from this approach do not remove the inherent differences between families. Fan et al. (2015) provide evidence from Norwegian registry data that the educational gender gap is correlated to mothers' labor force participation. They find evidence that boys are more susceptible at early ages if their mothers are employed. Kristoffersen et al. (2014) check whether gender differences in behavioral outcomes explain the gender gap in school outcomes. The authors find relationships between school outcomes and behavioral outcomes to be more sensitive to family and school environment for boys than girls. Buchmann and DiPrete (2006) explore how family background and academic achievement can affect the gender gap in educational outcomes. The authors provide evidence that boys are especially susceptible to the presence of a low-educated or absent father. Ferguson et al. (1994) document that family disruptions negatively affect children's test scores. They find no significant effect of parental separation on educational outcomes if separation occurs before school entry.

Allison and Frustenburger (1989) examine how children are affected by marital dissolution. The authors do not find evidence that boys are worse off than girls after marital dissolution. Owens (2016) documents that behavioral problems in early life can explain the gender gap in longer-term educational outcomes. Cherlin et al. (1991) use a descriptive study to show that much of the detrimental effect of divorce existed prior to separation. Interestingly these pre-divorce effects seem to be higher for boys than for girls. Kalil et al. (2015) use administrative data from Norway to show how father presence affects the intergenerational transmission of educational attainment. They find that more prolonged paternal exposure increases father-child association in education. Liu (2007) has used data from the panel study of income dynamics (PSID) to model educational attainment in a duration framework. The author has found evidence that educational attainment is affected through family disruption differently for girls and boys. This study is the first to use a sibling fixed effects model to study the gender gap in long-term educational outcomes.

This paper also contributes to the literature strand that analyses how family structure changes impact behavioral outcomes. Cobb-Clark and Tekin (2011) show that adolescent boys engage in more delinquent behavior if no father is present. They find that the girl child's tendency to show delinquent behavior is not affected by the father's presence. Prevoo and Ter Weel (2003) find statistically significant correlations between family disruptions before age 16 and personality development. The study finds that divorce has the most considerable negative effect of all the disruptions. This effect reduces with age and is stronger for girls. Bertrand and Pan (2013) thoroughly explore how family background affects gender differences in early childhood outcomes. The paper documents that boys raised in single-parent families have twice the behavioral and disciplinary issues rate as boys raised in two-parent families and are more than twice as likely to be suspended from school by the eighth grade.

Finally, this paper also adds to the literature looking at the impact of family disruption on long-term educational outcomes. Keith and Finlay (1988) use National Survey data to find that parental divorce leads to lower educational attainment and marriage age for both sexes. Interestingly they find that daughters of divorced parents have a higher probability of getting divorced. Krein and Beller (1988) document that the effect of living in a single-parent family increases with the number of years spent in that specific kind of family. This effect is most potent in preschool years and higher for sons than daughters. The authors have used matched Mother-Son and Mother-daughter samples from the National Longitudinal Surveys. Chetty et al. (2016) also examine the effects of childhood disadvantage on gender gaps in adult life. They find evidence that poor childhood living conditions lead to boys having lower formal employment rates. The authors define poor childhood living conditions using family income, neighborhood minority concentration, and crime rate. This analysis has been performed with the help of population tax records for children born in the 1980s. These results support my hypothesis that boys will be more negatively affected if their parents get divorced at a young age. I look at longer term outcomes such as high school graduation, college attendance[1], college graduation, crime[2], poor health[3] and idleness[4]. My results show

[1]The highest level of education reported in the last survey year is used as the educational outcome

[2]This variable is set equal to one if the respondent has ever been in jail, been convicted, been sentenced or been on probation.

[3]Since the 1994 wave, the NLSY has been asking respondents to describe their present health. They offer them five options- Excellent(5), Very Good(4), Good(3), Fair(2), and Poor(1). I code responses less than good on average as poor health.

[4]Respondents not earning a positive salary or enrolled in school are considered idle.

that when parents are divorced, the male child's educational outcomes are more likely to be negatively impacted.

2.3 Data

The primary data set I use in this analysis is the CNLSY. In 1986, the National Longitudinal Survey (NLS) began a separate survey of the children of the 6,283 women in the NLSY. The CNLSY tracks every child born to an NLSY respondent, enabling a comparison of siblings within the same family. Furthermore, mothers are surveyed extensively before the birth of their children, which allows for a rich set of controls for early life circumstances. I use the NLSY 1979 to get this information. The NLSY contains information on youths aged 14-22 in 1979. These youth are then surveyed using a detailed questionnaire every two years. This provides the date of marriage and the date of the end of the marriage. I restrict my analysis sample to children who are full biological siblings. One drawback of the CNLSY is that it is not straightforward to distinguish between full, half, or even adopted siblings. Hence, looking at divorce's impact in a within-sibling analysis is difficult. I identify biological siblings using Rodgers et al. (2016). The researchers provide a kinship link for respondents of the CNLSY. I use this to make the distinction between biological children and stepchildren.

The primary analysis sample has families with multiple children with at least one child from both genders. To check the impact of early divorce on the gender gap in long-term outcomes, children labeled as early divorced need to be less than 13 when their parents get divorced. The primary analysis sample is also restricted to respondents with a minimum age of 19. I remove anyone below 26 for the regression on the variable of college graduation and 22 for college attendance. I only consider children from married families in my analysis. The sample size for the primary regression is 2842 with 656 children from families where parents get divorced before age 13. In the comparison group, 633 children experienced a parental divorce after turning 13, and 1564 never experienced it. So, there are 656 children belonging to families with at least two children of different genders whose parents are divorced before they turn 13. I use these 656 children to identify the effect of early divorce on the gender gap in high school graduation. Figure 2 plots the distribution of children's age at the time of parent's divorce by gender. For this figure, I am considering children whose parents are divorced on or before they turn 18. There are 962 children whose parents have divorced before they become adults. Of the 962, there are 478 girls and 484 boys; these are the children I use to identify the effect of divorce from my exposure time and age group models. There is a slight spike in the number of girls whose parents get divorced while the children are in their early infancy. The pattern reported by the figure is similar to the results of Dahl and Moretti (2008) in which they show that families with firstborn daughters are much more likely to get divorced.

Table B.1 presents the summary statistics of the outcome variables by gender and family structure. Children are classified as late divorced if their parents are divorced after they turn 13 and never divorced if their parents have never been divorced. The gender gap in high school graduation, college attendance, and idleness deteriorate when moving from late divorced to early divorced families. The gender gap for all outcomes is considerably worse for boys when compared between divorced and never divorced families. The table reveals that girls from early divorced families are less likely to be idle than girls from late divorced families. Girls from early divorced families may be more likely to get divorced after marriage. This may put extra pressure on single mothers to find employment.

Figures 3 and 4 plot cross-race and income differences in the gender gap in educational outcomes between early and late divorced children. I am defining the children from families in the top quintile of permanent income[5] as high income and children from the bottom quintile as low income. The vertical axis lists the graduation/attendance rates, and the horizontal axis the educational outcomes by gender and family structure. I use the primary regression sample for this diagram. The figure provides visual evidence of the difference in gender gap between early divorced and late divorced families. The gender gap in high school

[5]Average family income before 18.

graduation increases for low-income families, while the gender gaps in college attendance and graduation go up for high-income families. When splitting the sample by race, there is a slight divergence in educational outcomes for children of both races, with big increases in college attendance for non-white and high school graduation for white children.

2.4 Empirical Strategy

I use a difference-in-differences (DD) style identification strategy with sibling fixed effects to see how parents' divorce during early childhood affects the gender gap in long-term labor market outcomes. The relationship between parents' divorce, child gender, and my outcomes of interest are formalized as follows:

$$Y_{ij} = \alpha + \beta_1 boy_i + \beta_2 boy_i * earlydivorce_i + \beta_3 earlydivorce_i + \beta_4 boy_i * white_j + \beta_5 X_i + \gamma_j + e_{ij} \tag{2.1}$$

$$Y_i = \alpha + \beta_1 boy_i + \beta_2 boy_i * earlydivorce_i + \beta_3 earlydivorce_i + \beta_4 boy_i * neverdivorce_i + \beta_5 neverdivorce_i + \beta_6 boy_i * white_i + \beta_7 white_i + \beta_8 X_i + e_i \tag{2.2}$$

The coefficient β_2 captures how parents' divorce before age 13 affects boys' outcomes differently than girls. Y_{ij} is the outcome of child i in family j. $earlydivorce_i$ is an indicator for child i's parents getting divorced before they turn 13. $neverdivorce_i$ is an indicator for child i's parents not getting divorced before the child turns 18 and is added to (2) due to the absence of the sibling fixed effects term γ_j. Control variables that differ across siblings and can affect the outcome, including birth order, birth month, and age of mother at birth,etc; are included in X_i[6]. The impact of a child's race may vary by gender and is included as an interaction term with the indicator boy_i.

The primary assumption which helps interpret the coefficient β_2 as the causal effect of early divorce on the gender gap is that the latent gender gap in sibling outcomes is

[6]Age difference and birth order are both important within sibling variations. The mother's age at birth is perfectly correlated with the relative age gap between siblings.

independent of divorce before age 13. I test the plausibility of this assumption by running multiple placebo tests on achievement tests administered by the CNLSY. I also provide event study analyses that show clear visual breaks from the trend in achievement scores right around the time of the divorce. This also helps to show that there are no pretrends in my analysis. The definition of early divorce helps to identify whether a boy will be more harmfully impacted than his sister if he is less than 13 at the time of divorce. While the 13-year-old cutoff may be conservative, it also minimizes the probability that some children have left home. I am also presenting results for running the same regression mentioned above without the sibling fixed effects, i.e., simple OLS. I am using the same sample for the OLS and sibling fixed effects models.

2.5 Results

The educational outcomes I look at are high school graduation, college attendance, and college graduation. Additionally, I look at three different non-educational outcomes. First, I construct a dummy variable called crime. This variable indicates whether the respondent has been in jail, sentenced, convicted, or on probation. Second, I define respondents not earning a positive salary or enrolled in school as idle. Third, since the 1994 wave, the NLSY has been asking respondents to describe their present health. They offer them five options- Excellent(5), Very Good(4), Good(3), Fair(2), and Poor(1). I code responses less than good on average as poor health.

2.5.1 Educational Outcomes

The second column of Table B.2 reports the regression estimates for the high school graduation outcome with sibling fixed effects. The coefficient β_2 tells us that boys' likelihood of graduating high school relative to their sisters goes down by 6.4 percentage points due to early divorce. Boys' are also 15.6 percentage points less likely to graduate high school than their sisters if their parents had been divorced before they turn 13. Relative to baseline estimates (The average high school graduation rate in my entire sample is 84.4 percent), this translates to an 18 percent decrease in high school graduation. This result carries significant economic consequences. Males under age of 40 with high school or lower education have seen their real earnings drop by 25 percent between 1979-2010 (Autor and Wasserman 2013). Column 4 looks at the effect of early divorce on the gender gap in college attendance. Early divorce reduces the likelihood of boys attending college relative to their sisters by 8 percentage points. Boys from early divorced families thus have a 23.9 percentage point lower chance of attending college than their sisters. Relative to baseline estimates[7] this translates to a 40.8 (23.9/58.5*100) percent decrease in college attendance.

The estimates for the effect of early divorce on the gender gap in college graduation are neither economically nor statistically significant. The odd numbered columns in table B.2 provide regression results with the fixed effects sample but by running just OLS. I notice that the coefficient on β_2 becomes more negative as we move from OLS to the fixed effects models. This may be due to the fixed effects model taking care of family-specific variation. I have provided standardized values for the educational outcome results in Table B.9.

2.5.2 Non Educational Outcomes

Table B.3 presents estimates for non-educational long-term outcomes using equation 1. I find no evidence of a strong impact of early parental divorce on the gender gap in non-educational outcomes. Column (6) reports an increase in the likelihood that girls self-report poor health compared to their brothers by 3 percentage points. However, the small effect size suggests a modest impact at best. Boys from early divorced families are more likely to commit a crime by 1.3 percentage points relative to their sisters. Divorce before age 13 increases boys' likelihood of being idle compared to their sisters by 2.4 percentage points. The effect sizes for all three outcomes are almost negligible and do not point to a strong impact of parental divorce.

[7] The average college attendance rate in my entire sample is 59 percent.

2.5.3 Educational Outcomes By Demographics

Table B.4 reports the primary regression outcomes by race (white and non-white). Column 1 of Table B.4 shows that early divorce strongly impacts white children's gender gap in high school graduation. The estimates show that early divorce reduces the likelihood that a white boy will graduate high school by 10.6 percentage points relative to his sister. White boys from early divorced families have a 16.8 percentage point lower chance of graduating high school than their sisters. As shown in columns (3) & (4), the difference in magnitude of the estimates between races is much lower for college attendance. These results also have higher standard errors due to smaller sample sizes, leading to imprecise estimates. Columns (5) & (6) report a large gap in the magnitude of estimates for college graduation. Like columns (3) & (4), these results are noisy due to large standard errors.

Table B.5 reports the primary regression outcomes by income level. I consider only the top (which I define as high-income) quintile (top 20 percentile) and bottom (low-income) quintile of permanent income in my sample. Early divorce does not have a statistically significant effect on the gender gap in educational outcomes after splitting the sample by income. Children from families in the middle of the income distribution may be driving the results using the entire sample. However, early divorce does seem to have a strong negative effect on the gender gap in children's outcomes within the same family for college graduation in high-income families and college attendance in low-income families. Nonetheless, these results are imprecise due to a lack of sufficient power.

2.6 Exposure Time Models

The regression specification in equation 1 uses a particular cutoff age to try and measure the effect of divorce on the gender gap in children's outcomes. The early divorce variable in the sibling fixed effects model looks at whether a child was less than 13 versus greater than 12 at the time of divorce. This approach measures the effect of the timing of the divorce on the child. It may be the case that boys are worse off if parents get divorced at an earlier age than girls. It may also be that girls are less susceptible if their parents get divorced when they are teenagers. Thus, the timing and intensity of divorce vary across children in the same family. To overcome this hurdle, I run exposure time models. I have run two models, with and without fixed effects, on the same sample used for the previous regressions. There are 962 children in the primary regression sample whose parents are divorced before they turn 18. They consist of 484 boys and 478 girls. These 962 children are used to identify the effect of parents' divorce during childhood on the gender gap in outcomes for the exposure time models.

$$Y_{ij} = \alpha + \beta_1 boy_i + \beta_2 boy_i * exposure_i + \beta_3 exposure_i + \beta_4 boy_i * white_j + \beta_5 X_i + \gamma_j + e_{ij} \tag{2.3}$$

$$Y_{ij} = \alpha + \beta_1 boy_i + \beta_2 boy_i * young_i + \beta_3 boy_i * midchild_i + \beta_4 boy_i * teenager_i + \beta_5 young_i + \beta_6 midchild_i + \beta_7 teenager_i + \beta_8 boy_i * white_j + \beta_9 X_i + \gamma_j + e_i \tag{2.4}$$

The first uses a linear function of (18-age) at divorce in place of the divorce dummy. This variable is defined as exposure in the equation. The variable of interest β_3 measures the effect of an additional year post-divorce through age 18 in boys relative to girls. Suppose two children are in a family, and their parents get divorced before either turns 13. The divorce cutoff variable will be one for both. However, both children may have different lengths of exposure to the divorce. This specification utilizes this variation in the timing and intensity of divorce. For the second model, I am using age groupings. I am using three age groupings in the main paper. They are young(ages 0-4), midchild(ages 5-12) and teenager(ages 13-18). The impact of the length of exposure to the divorce may not be linear. Also, family disruptions at different points of time may have varying impacts on different outcomes. For example, parental divorce during early childhood may be particularly harmful for high school graduation, but divorce during teenage years is more harmful for college graduation. I now have three variables of interest β_1, β_2 and β_3. Each captures the differential impact of being divorced in a particular age group on boys relative to girls. The results of the age group regressions are robust to choosing adjacent years in a three-group setting or increasing the number of groups to four. These models are run for both the sibling fixed effects and standard OLS.

Table B.6 reports the results for the first exposure time model. Column 4 tells us that a one-year rise in exposure to divorce reduces the likelihood of attending college by 0.6 percentage points for boys. The average male child from a divorced family in the analysis sample is exposed to divorce for eight years before turning 18. Hence, on average, boys from divorced families have a 20.7(0.6*8 + 15.9) percentage point lower chance of attending college than their sisters.

Table B.7 reports estimates using equation (4). Column 2 looks at the outcome of high school graduation. Parents getting divorced in the middle of a boy's childhood increases the within-family boy-girl disparity in high school graduation by 7.4 percentage points. Divorce during teenage years negatively impacts boys' likelihood of attending college relative to their sisters. Column 4 reveals that boys impacted by divorce in their teenage years are 12.2 percentage points less likely to attend college than their sisters. The estimates from column 6 do not allow precise measurement of the impact of divorce on college graduation.

2.7 Placebo Tests

Interpreting the above estimates as the causal effect of early divorce on the gender gap in child outcomes relies on the assumption that the latent gender gap in sibling outcomes is independent of future marital status. Falsification tests are provided by analyzing the relationship between early divorce and the gender gap in achievement scores for children whose parents are yet to be divorced. Outcome variables used in the falsification tests are Peabody Individual Assessment Math (PIATMT) subtest, Reading (PIATRR) subtest, Behavioral Problems Index (BPI) score, and Home inventory score (HOME) for ages 5 and 6. The sample consists of children whose parents are either never divorced or divorced after the children turn seven. The outcome variable for each regression is the percentile score for that particular test.

The results in table B.8 support the claim that there is no inherent difference in the gender gap in educational outcomes prior to divorce. While there is some imprecision in the estimates, little indicates that early divorcing families have a different gender gap from late divorcing families. Table B.10 provides standardized values for the placebo test results.

2.7.1 Attrition Effects

The CNLSY takes interviews of its respondents every two years from the 1986 wave. Several respondents may not show up in every survey wave for several reasons, including incarceration, military assignment, major illness, leaving the parent's household, or changes in survey cutoffs for young adults (between waves). CNLSY interviews these respondents inconsistently. For example, in several waves, some respondents are interviewed while in jail. Two survey waves have different young adult cutoff ages; interviews for adult respondents over 30 occur every other wave. These attrition effects may sometimes bias the estimates for the long-term outcomes. The primary regression models are run on two samples to check whether this causes a bias in the estimates. First is the primary sample without incarcerated respondents. Second is the primary sample with respondents who appear in at least 80% of the waves. There are no significant changes in the outcomes for any of these regressions. This robustness check further clarifies that no significant attrition effects are biasing the estimates. Tables 11 and 12 provide results for the non-incarcerated and consistent response samples.

2.7.2 Alternate cutoffs

These results are not sensitive to different age cutoffs. While using adjacent ages like 12 and 14 yield very similar results (as expected), I have also checked the sensitivity of these results at ages much farther apart. Tables 13 and 14 provide results using age cutoffs between five and 15.

2.8 Event Study Models

I use the timing of divorce to carry out an event study analysis of achievement scores. This regression model assesses whether divorce leads to an impact change in trends in achievement scores. I provide event study analysis for the fixed effects model. The event studies are generalized difference-in-differences models similar to Jacobson et al. (1993). The event study regression equation is below:

$$Y_{ijt} = \alpha + \beta_1 boy_i + \Sigma_{\tau=-3}^{-2}\delta_\tau Div_{\tau i} + \Sigma_{\tau=0}^{3}\delta_\tau Div_{\tau i} + \Sigma_{\tau=-3}^{-2}\phi_\tau Div_{\tau i} * boy_i + \Sigma_{\tau=0}^{3}\phi_\tau Div_{\tau i} * boy_i + \beta_2 boy_i * white_j + \beta_3 X_i + \gamma_j + \epsilon_{ijt} \tag{2.5}$$

Each period in the event study analyses equals two years with a maximum of three periods to account for the biennial nature of the CNLSY, and the limited number of school-age children administered the achievement scores by the CNLSY. I provide event study diagrams for childhood math and reading achievement scores for the sibling fixed effects model. Figure 5 reveals that the gender gap in math assessment scores remains almost constant till the time of divorce, after which there is a sharp fall. There is a similarly sharp fall for reading assessment scores in figure 6, where the largest drop happens in the period before the divorce. There is no presence of any pre-trends in the math test scores. These figures provide visual evidence supporting the absence of parallel trends.

2.9 Conclusion

This paper investigates whether divorce exerts a differential effect on the labor market outcomes of boys relative to girls. I use the National Longitudinal Survey of Youth 1979 Child & Young Adult Supplement (CNLSY) and a model that employs within-family brother-sister comparisons. I find that divorce disproportionately negatively affects the long-term educational outcomes of boys relative to girls. The event has severe consequences for boys whose parents are divorced while they are teenagers. Divorce during the teenage years increases the boy-girl disparity in attending college by 14.4 percentage points. This translates to a 39.7 percent lower likelihood of attending college. Divorce before age 13 reduces a boy's likelihood of graduating from high school by 15.5 percent relative to his sister. I find no significant effect of divorce differing by gender on non-educational long-term outcomes. My findings are robust to different specifications, including an age cutoff and two different exposure time models. I estimate event study models and placebo tests in reading and math scores to check for the validity of my assumptions. The event study models and placebo tests provide evidence for the absence of pretrends and independence of future marital status to the gender gap in achievement scores, respectively.

I find that the boy-girl disparity in high school graduation increases if parental divorce happens when the child is less than 13, especially when the child is between 5 and 12. I find that divorce during the teenage years significantly affects the gender gap in college graduation. Parental divorce during latter ages may have more potent effects on more advanced educational outcomes.

There are multiple channels through which divorce may affect boys differently than girls. There is a possibility that the dissolution of a turbulent marriage may have a less harmful or even positive impact on the male child. It may also be the case that the absence of a father figure with a negative influence may positively impact a boy's life. However, I do not find any results that may support these hypotheses. My analysis provides evidence of different channels through which divorce may be affecting children. First, boys may be more responsive to changes in family structure than girls. Second, with changes in family structure, the relative time invested in boys may start to fall with age compared to girls. These hypotheses are supported by Baker and Milligan (2013), Bertrand and Pan (2013), and Bibler (2019). The data in the CNLSY is not sufficient to test these channels.

Policymakers must be careful in trying to solve this problem. While it is the case that families with male children are less likely to go through a divorce (Dahl and Moretti 2008), the consequences of divorce are also more severe for boys. It may be more fruitful to find effective ways to target boys from divorced families who are most at risk and better understand the reasons behind broken marriages of families with children. Policymakers should target the children suffering from this unfortunate accident and find a way to reduce such events in the future. Any progress can not only reduce some unfortunate suffering but also increase the pool of productive adults in society and reduce the already propagating gender disparities.

CHAPTER 3

HOUSING WEALTH AND PRIVATE SCHOOL ENROLMENT

3.1 Introduction

A long literature in economics has focused on parents' school choices for their children. Becker (1960) introduced children into economic models as a durable good in the utility function of the parents. Because parents also derive utility from the quality of the child, education and future earnings potential become essential factors. Families have become more aware of the remarkably high recent labor market payoffs to educational attainment. Private school tuition has also risen over the past few decades. Private schooling may be another way high-income families seek to give their children an advantage in preparing for post-secondary education. With increasing wealth and income, relatively affluent families can afford the high tuition that most nonsectarian and some religious private schools charge. They may also increasingly want to substitute payment of private school tuition for the time they would otherwise spend monitoring their child's experiences in public schools.

I estimate multinomial logit models among homeowners of the likelihood of switching from a private school to public school, staying in private school, or switching from public school to private school, with staying in public school as the omitted category, as a function of home price growth. I also control for a detailed set of background characteristics, including religion and household income. The empirical strategy is to compare the schooling choice of students who experience varying magnitudes of housing price changes at different points of their K-12 education career.

I find that home price variation affects school choice. A \$100,000 increase in home prices increases the relative probability of switching from private to public school by 0.007 percentage points. When looking at estimates by different groups, I find the strongest effects of switching from private to public school for families in the first income quartile and Non-

Catholic families. I also find a strong effect of house prices on the relative likelihood of switching from public to private school when children transition from middle to high school. A $100,000 increase in home prices increases the relative probability of switching from public to private school by 0.009 percentage points.

The recent boom and bust in housing prices has received much interest from economists. Housing wealth changes are perceived as permanent changes in wealth and affect current consumption. Many studies have found a positive relationship between housing wealth and consumption (Hurst and Stafford 2004; Lehnert 2004; Case, Quigley and Shiller 2005; Campbell and Cocco 2007). Studies have shown housing wealth to affect a wide variety of outcomes. They include college enrolment (Lovenheim 2010), college choice (Lovenheim and Reynolds 2011), fertility (Lovenheim and Mumford 2013) , labor supply (Zhao and Burge 2017), etc. Currently, little is known about how house prices affect private school enrolment. Given that less than 10 percent of American children attend a private elementary or secondary school, why should we care if housing wealth impacts private school enrolment? Suppose the private schools provide a better education than those available to children from lower-income families. In that case, these choices pass the economic advantage to the next generation and undercut the potential for intergenerational economic mobility.

Understanding the effect of housing wealth on private school enrollment is crucial from various other perspectives. First, private schools are an integral part of our K-12 education system, so it is vital to understand whether and how housing wealth affects private school enrollment. Moreover, anything that affects enrollment in these schools has the potential to have significant consequences on school quality and the educational outcomes of students (both public and private). Second, an important factor is a potential impact on per pupil spending in public schools.

Suppose it is indeed the case that a significant number of public school students are now transferring to private schools. In that case, this may increase the per pupil spending in public schools unless total school spending decreases at a corresponding rate. On the other hand, fewer children in the public sector decrease the number of people with stakes in the quality and performance of public schools. This, in turn, may reduce demands for more resources in the public sector. Thus, while the exact direction of the net effect on funding is unclear, housing wealth impacts private school enrollment and can impact per pupil funding in public schools.

In this paper, I use family wealth variation supplied by the housing market to identify how household resources affect children's schooling decisions. My analysis makes several contributions to the existing literature. This is one of the few studies to examine how children's school choice responds to the household's wealth rather than simply its income. Excluding wealth may be particularly problematic because it can cause one to mischaracterize the financial resources of the household. If school choice decisions are a function of total resources, using income as a proxy may yield an incomplete picture of how resources affect private school enrolment. I analyze housing wealth as a measure of household wealth for several reasons. First, about 60 percent of families of the Child and Young adult supplement of the National Longitudinal Survey of Youth(NLSY), the data I use in this analysis, are homeowners. Second, for these families, and the United States as a whole, housing wealth represents the vast majority of total household wealth[1]. Third, about 77% of private school attendees come from families with a home. The main identifying variation in this analysis comes from the housing boom that began in the late 1990s and was characterized by large increases in home prices that occurred differentially across cities and increased liquidity of home equity. Home owners who lived in high-growth areas experienced a large increase in their liquid wealth relative to home owners in lower-growth areas and renters throughout the United States. In examining the effect of housing wealth changes on school choice decisions, my paper thus contributes to the growing literature on housing wealth and household behavior.

Whether family resources affect k-12 education investment decisions is an important em-

[1]For example, in the 2004 survey of consumer finances, home equity accounted for 85% of household wealth for the median homeowner between 25 and 55 years old.

pirical question in economics. The relevance of this question is underscored by the large differences across the income distribution in private school enrolment. For example, in my data, the private school enrolment rate among children from families in the lowest third of the income distribution is 5 percent, while among those in the highest third is 20 percent. As the differences in private school enrolment across the income distribution may not indicate a causal role for family finances, it is important to identify whether family resources have a causal effect on the decision to invest in private education. If families use income and wealth to pay for private school, excluding wealth will cause one to mismeasure the empirical relevance of household finances. Identifying wealth's causal effect on private school enrollment is difficult even if wealth measures are available. This is because families that accumulate wealth are typically more likely to send their children to private school due to unobservable attributes that correlate with savings behavior and education, such as preferences for private school education.

Housing price changes may affect private school enrollment through other channels. Between 1990 and 2005, the real median home price in the United States rose by 55%, and extracted home equity as a percentage of personal income rose by 600%. This can potentially impact credit-constrained students whose families do not have sufficiently affordable access to credit. When homes go up in value, it increases the amount of equity that can be used in collateral. Home equity loans are relatively inexpensive; thus, home price changes can reduce the interest rates at which a family can borrow funds for private school education. Hence, families that experience large increases in home price would have a significantly easier time financing expenditures due to the increased ease of borrowing against their home's value (Bennett,et al 2001; Deep and Domanski 2002; Greenspan and Kennedy 2005; Doms and Krainer 2007).

Parents may consider children's private school education as a consumption good. In this case, housing wealth variation can affect private school enrollment via a direct wealth effect. Rising home prices make homeowners wealthier. This increases the likelihood of homeowners consuming more goods. Private school education for their children may be one of these goods. Finally, higher home prices can increase school revenues via local property taxes. This can have a positive impact on school quality. This may increase the appeal of public schools to other parents leading to switching away from private schools to public schools. Similarly, choosing to send their children to private school may also increase the opportunity cost of losing out on the public school subsidy.

The empirical strategy and data I am using for this study make it difficult to separate the relative importance of the underlying mechanisms. Given the large fluctuations in home prices and the high value of K-12 education, studying the effects of home price changes on private school enrolment has significant policy relevance, even without the underlying mechanisms for effect. The rest of the paper is organized as follows: Section 2 reviews the literature. Section 3 describes the data. Section 4 presents the empirical strategy. Section 5 presents the estimation results. Section 6 concludes.

3.2 Literature Review

There is an active literature studying the effect of housing prices on educational outcomes. Lovenheim and Reynolds (2010) use the NLSY-97 to study the impact of housing wealth changes on college choice. The authors provide evidence that increased housing wealth increases the likelihood of attending a public flagship college relative to a non-flagship college and decreases the likelihood of attending community college. This individual level analysis is helped by the level of detail in the NLSY. Lovenheim (2011) looks at how liquid housing wealth affects college enrollment. The study shows that short-run changes in housing wealth affect college enrollment. The effect is most substantial for low-income families. The author uses the PSID to construct a repeated cross section of 18-19-year-olds. Charles et al (2018) provide a detailed analysis of how the boom and bust in housing that lasted from the late 1990s to the late 2000s affected college enrollment. The study is one of the first to examine the effect of both the housing boom and the housing bust. The study provides evidence at both the individual and aggregate levels. The housing boom and bust have opposite effects on college enrolment. They analyze data from the Census, integrated post-secondary education data system (IPEDS), federal housing finance agency (FHFA), and NLSY-97. The authors estimate structural breaks in the evolution of housing prices at the MSA level. They use the variation in timing and intensity of changes in housing prices to conduct an event study analysis of employment and college enrolment. The authors argue that all else equal, the housing boom should most affect the students on the margin of going to college at all (community college) and have little effect on investment in Bachelors-level training. To my knowledge, there has been no study on the effect of housing wealth on private school enrolment.

There is an emerging body of work studying the effects of housing wealth changes on outcomes other than education. Lovenheim and Mumford (2013) use data from the panel study of income dynamics (PSID) to explore how housing wealth affects fertility. The authors use PSID restricted-use microdata from 1985-2007. The study finds that a 100,000 dollar increase in housing wealth causes a 16 to 18 percent increase in the probability of having a child. The authors also provide results showing increased responsiveness of fertility to housing wealth over time. Zhao and Burge (2017) use health and retirement survey data to explore the relationship between housing wealth and labor supply. The authors exploit the exogenous variation in housing wealth to use a within-MSA renter versus home owner difference in difference approach. They find evidence that housing prices and property taxes affect labor supply in opposite directions. Benjamin et al (2004) compare the effects of real estate and financial wealth on consumption. They use aggregate consumption and wealth data from the National Income and Product Accounts(NIPA) and the Federal reserves flow of funds(FOF) accounts for the years 1952-2001. The authors find that the decline in consumption from the stock market crash of the dot-com bubble was majorly offset by increased real estate wealth. The study also shows that holding real estate wealth serves as a way of reducing volatility by consumption smoothing. Campbell and Cocco (2007) use

UK micro-level data to investigate how household consumption responds to housing prices. The study finds older home owners to have the largest house price elasticity. The authors also control for economy-wide house prices and regional income to find that regional house prices affect regional consumption.

The economics of education literature has seen a steady flow in the number of papers studying private schools. The existing studies primarily focus on the impact of other schooling options on private school enrolment. There is also some literature on how race and background affect private school enrolment. There is very little work on how housing wealth affects private school enrolment. Fairlie and Resch (2002) and Li (2009) examine white flight into private schools. Both studies find evidence of white flight to private schools. Fairlie and Resch use national education longitudinal study (NELS) data to run probit regressions for the probability of attending private schools. The authors find negative estimates of the effect of minority share on switching from private to public school in the eighth grade. Li (2009) uses data from the High School and Beyond (HSB) to show that white parents are more likely to send their children to private schools in areas of high minority student concentration. Poor minority students have a stronger impact than non-poor minority students. The study finds that Blacks have the strongest effect among all minorities. This effect also differs by region, as the effect is less prominent in the West than in other parts of the United States. The author provides estimates at both the county and the Metropolitan area(MA) level.

Betts and Fairlie (2003) investigate whether immigration may have induced native flight among native whites. The authors use data from the Census on 132 metropolitan areas and a quasi-difference-indifference approach in their investigation. The study finds significant effects for secondary school enrolment and no strong effects for primary school enrolment. The impact on white natives is strongest from immigrant children who do not speak English at home. A recent study by Chakraborty and Roy (2016) examines the impact of charter schools on private school enrolment. The authors use data from the national center for education statistics (NCES) biennial private school surveys. The study exploits the exogenous variation created by the Michigan charter law, along with a fixed effects and instrumental variable strategy to investigate the effects of charter penetration on private school enrolment. The authors do not find any significant effect of charter penetration on private school enrolment. The results do not change after splitting the sample by religious affiliation. This proves that recent private school closures were primarily driven by increasing costs, tuition, and teacher salaries. This study may not be representative of the entire country as it is restricted to the state of Michigan.

Barrow (2006) finds that private schools are located in areas of higher income dispersion and racial diversity. The study also finds evidence that private schools are less likely to be located in areas of high-income homogeneity and racial concentration. Murnane and Reardon (2018) use multiple nationally representative data sets to study long-term enrolment trends in private schools. Like previous studies, income inequality is a major factor in private school enrolment. Affluent families who live in cities are much more likely to send their children to private schools than those who live in the suburbs. Families in the South and West are more likely to send their children to private schools. The 90-50 gap, the difference in private school enrolment between families at the 90th percentile and median of the income distribution, has increased considerably in the past few decades. Lower no. of church staff involvement in Catholic schools has led to increased operating costs. This has caused many Catholic schools to shut down in the past few decades.

Buddin et al. (1998) examine the factors driving parents' choice of private school. To estimate their model, the authors use data from the 1990 census and supplemental information on California's private and public schools. The authors find taste and income to be major factors driving parents' choices. They find that private school choice is not sensitive to private school costs. Long and Toma (1988) also find similar results. Private school attendance was determined by the usual factors such as income, race, religion, and supply of private schools. They provide evidence that the difference in enrolment by income and race narrowed in the 1970s. This has changed drastically in recent papers. Pandey et al. (2009) use data from the National Center for Education Statistics(NCES) to examine private school closures. This is one of the few studies in the literature looking at the supply of k-12 private schools. The study shows that most non-profit schools are market driven. The study also examines the relationship between private schools and public school quality. School districts with high student-teacher ratios have lower private school failure rates. The authors model private school closure on the excess supply of private schools over their demand and the individual school's characteristics. Goldring and Philips (2008) explore the dynamics surrounding school choice. The study uses school choice survey data from Metropolitan Nashville public schools. The authors try and find demographic factors driving parents' private school choices. This study also looks at other channels: parental involvement, satisfaction, priorities, and social networks.

Catholic schools constitute a large share of private schools in the United States. A vast amount of literature in economics studies the effects of Catholic schools on educational outcomes and the factors explaining Catholic school enrolment. Elder and Jepsen(2013) explore whether Catholic schools are more effective than public primary schools. They find Catholic schools to have significantly better average outcomes than public primary schools. However, they also provide evidence that the advantages of Catholic schools are driven entirely by selection bias. Cohen-Zada and Elder (2009) show that county-level Catholic shares measured at the end of the 19th century are more strongly associated with Catholic school attendance than current Catholic shares. Altonji, Elder, and Taber (2005) show that Catholic school attendance is associated with a higher probability of graduating high school and attending college. Jepsen (2003) finds that Catholic primary schooling does not significantly affect mathematics and reading test scores. Rearden et al. (2009) find that Catholic schools are less successful at teaching math skills than public schools and are no different at teaching reading skills than public schools.

3.3 Data

The microdata I use for the present analysis is the public-use data from the CNLSY. In 1986 the National Longitudinal Survey (NLS) began a separate survey of the children of the 6,283 women in the NLSY (National Longitudinal Survey of Youth). The CNLSY tracks every child born to an NLSY respondent. Furthermore, mothers are surveyed extensively prior to the birth of their children. The NLSY contains information on youths aged 14-22 in 1979. These youth are then surveyed using a detailed questionnaire every two years. The main advantage of the NLSY over other available survey data is that it is a long panel that allows me to track changes in the family's reported home price during a child's school years. The data also contains a rich set of individual and family background information. I use the reported market value of the home as my home price measure. Home ownership status is defined as owning a home in the survey year.

Table C.1 presents summary statistics of the NLSY data, separately for home owners and renters. On average, 14% of home owners send their children to private school, compared to 5% of renters. Children born in home owning families are more likely to have older and more educated mothers. Home owners are also less likely to be racial minorities than renters. The average 4-year change in housing value for home owners is $44,000. Finally, home owners are higher income households and have a much smaller share of non-native English speaking families than renters.

3.4 Empirical Strategy

In order to test whether home price changes affects private school enrollment, I estimate multinomial logit models of the following form:

$$P(j^*_{ic} = j_{ic}) = \beta_0 + \beta_1 \Delta P^h_{ic} + \gamma X_i + \phi_c + \epsilon_{ic} \tag{3.1}$$

where i indexes individual and c indexes cohort. The variable ΔP^h_{ic} is the reported home price change in the relevant time period of the respondent's school going years. The vector

X is comprised of individual and family background characteristics.

This model is used in a cross-sectional setting. The model estimates the effect of housing wealth changes on changes in school type at different points in the childs education. I look at two crucial transition periods in a child's k-12 education, primary school to middle school and middle school to high school. ΔP_{ic}^{h} for these two models measures the change in self reported home price over this transition period. I also look at the change over the child's entire k-12 education, with ΔP_{ic}^{h} measured on the corresponding time frame. The four categories for the multinomial logit include starting in public and ending in public(baseline), starting in private-ending in public, starting in private-ending in private, and starting in public-ending in private.

I focus on recent individual-level changes because contemporaneous home price levels may be loosely related to household resources: families can own an expensive home without having any equity. However, individual home price changes are capitalized into housing wealth, making home price changes a more appropriate measure of wealth than the home price level. The identification assumption underlying equation (2) is that housing price changes are conditionally exogenous to the schooling decision. In other words, apart from the fact that housing prices increase household wealth, home price changes and private school enrollment should be uncorrelated conditional on the observables in the model. There are several threats to this assumption. A positive correlation between housing prices and local macroeconomic conditions may result in my housing price change measure picking up this relationship rather than identifying the effect of housing wealth changes on school enrolment decisions. Another potential threat to identification is the selection of households across areas. If parents planning to send children to private schools purchase homes in places most likely to experience high housing price growth shortly, my estimates will be biased upward. My data's lack of geographic information prevents me from tackling these problems. Finally, families may be adding significant upgrades to their homes. These upgrades usually result in increases in house prices. This could potentially threaten the validity of my estimates.

However, I do not have information on this question and cannot address this challenge.

3.5 Results

Tables 2 to 8 show results from different estimations from equation (1). Each table presents results from a single regression. Due to a lack of geographical indicators, the standard errors for the reported estimates are not appropriately clustered. To account for this, I provide estimates with robust standard errors and standard errors clustered at the region level. Accurate standard error clustering would have been done at the county or state level. Hence, the robust standard errors are likely to be lower than an accurate measure of standard errors. Similarly, the standard errors clustered at the region level may be too large. The true measure of standard errors is likely to be somewhere in between the two measures.

The columns of table C.2 show the marginal effects from multinomial logit estimates of housing price changes on the change in school relative to staying in public school. I find a strong relationship between home value change and switching from private to public school. A $100,000 increase in reported home prices increases the likelihood that a respondent switches from private to public school by 0.007 percentage points. Higher house prices may result in improvements in public school quality caused by rising school revenues. The relative importance of increases in public school quality may be higher than the wealth effect and possible reductions in credit constraints. Furthermore, I do not find statistically significant effects of house price changes on staying in private school or switching from public to private school.

The average effects in Table C.2 may mask heterogeneity across religions since Catholic schools play a major part in private education in the United States. To examine differences in the effect of home prices by Catholicity, I estimate a modified version of Equation 3.1 in which I interact home price changes with indicators for Catholicity. I find increases in the likelihood of switching from private to public school for Catholic and Non-Catholic children. I find that a $100,000 increase in home prices increases the likelihood of switching from private to public school by 0.009 percentage points for non-Catholic children and by 0.004 percentage points for Catholic children. Similar to results shown in Table C.2, increased school quality from higher public school revenues may be playing a large role in parents switching their children from private to public school with increases in home prices. Results presented in this table show that it is consistent across Catholicity. Home price increases cause larger increases in the likelihood of staying in private school for Non-Catholic children compared to Catholic children. However, all estimates apart from those reporting the likelihood of switching from private to public school are statistically insignificant.

To examine differences in the effect of home prices by family income, I estimate Equation 3.1 with interactions of home price changes with income group indicators. Table C.4 shows these estimates. I find large effects of switching from private to public school for the 1st and 4th income quartile. I find that a $100,000 increase in reported home price increases the likelihood of switching to public schools from private schools for children of parents in the 4th income quartile by 0.006 percentage points. High-income households likely live in an area with high home values. Sharp rises in housing prices may cause significant improvements in school quality via greater school revenues. This may be the primary factor driving wealthy parents to switch from public to private schools. For homeowners in the first income quartile, strong rises in housing prices can also mean significant improvements in the local public school quality. I find that a $100,000 increase in reported home price reduces the likelihood of switching to private schools from public schools by 0.014 percentage points for children from families in the second income quartile. I do not have any additional information to determine why households in this income quartile would reduce their likelihood of switching to private school. Additionally, I find substantial but statistically insignificant effects of house price changes on staying in private school for parents of children in the 1st and 3rd income quartile. The appeal of improved school quality through increased school revenue is likely the dominant cause behind my results.

Families may be more inclined to send children of a particular gender to private school.

I provide multinomial logit estimates by gender to test this hypothesis in Table C.5. I find large and statistically significant effects of switching from private to public school for male children. The probability that a student switches from private to public school increases by 0.008 percentage points for male children. Interestingly, there is a large difference between the estimates for switching from public to private schools between genders. Parents of female children are likelier to switch their children from public to private school than male children. For every $100,000 increase in home prices, the probability that a female child is switched from public to private school increases by 0.003 percentage points. Parents may believe that girls are more likely to benefit from private education. I do not have information to test this hypothesis.

Table C.6 estimates the relationship between housing wealth and private school enrollment by race. I find large and statistically significant effects on the likelihood of switching from private to public school for Hispanic children. For every $100,000 increase in home prices, the probability that a Hispanic child switches from private to public school increases by 0.008 percentage points. The standard errors for Black children are much higher than those of the other two races, leading to less precise estimates. Black children are also much less likely to switch from public to private schools. For every $100,000 increase in home prices, the probability that a Black child is switched from public to private school decreases by 0.007 percentage points. The table reports large effect sizes across all three races for switching from private to public school. These results show evidence that the appeal of improved public school quality through increased school revenue is high across all races.

In general, there are three stages of education post kindergarten in the United States. These are roughly: primary school (grades 1-4), middle school (grades 5-8), and high school (grades 9-12). Some major differences between these stages are campus size, the number of students in each class, the accessibility of teachers, how lessons are implemented, student expectations, and the interaction with families. Since these are natural break points, parents may be more inclined to move children to a private school at these junctures.

In Table C.7, I look at the impact of housing price changes on the likelihood of switching to a different school sector when transitioning from middle to high school. I find positive but mostly statistically insignificant effects of an increase in home prices on change in the school sector during the transition from middle to high school. I also find a small, statistically insignificant effect of house price changes on the likelihood of staying in private schools. Reported house price changes have the strongest effect on shifting from public to private schools. A $100,000 increase in home prices increases the likelihood of switching from public to private school during the transition from middle to high school by 0.009 percentage points. The transition to high school is a major event in a child's educational career. High school performance is a strong determinant of college attendance and college quality. These factors may influence parents' decisions by increasing the relative importance of the wealth effect or a less binding credit constraint.

Table C.8 shows results for transitions from primary to middle school. During this transition period, the largest effect sizes are for switching private to public schools. A $100,000 increase in home prices increases the likelihood of switching from private to public school during the transition from primary to middle school by 0.006 percentage points. Most estimates for the effect of house price changes during this transition are statistically insignificant. Like most of the earlier tables, increased school quality through increased school revenues may be the primary reason behind the increased likelihood of switching from public to private schools. Parents may also be less willing to invest in private education when switching from primary to middle school.

According to the NCES, prekindergarten, elementary, and middle private schools have decreased from 21,611 in 2009-2010 to 18,870 in 2019-2020. However, private secondary and high schools have increased from 3,405 in 2009-2010 to 3,626 in 2019-2020. The increase in private high schools may indicate a greater demand for private education among parents of high school children.

This study has several limitations. The data set used in this analysis does not have information on the respondent's location. This only allows for using changes in self-reported housing prices. I can also not identify respondents who are moving during the study period. Lack of location information also prevents me from controlling for local public school quality. Parents' decisions to send their children to private school may also be affected by local macroeconomic conditions and not changes in housing wealth. I am unable to separate this effect without location information. The primary identifying assumption in the specification is that the background characteristics are sufficient to control for any selection of families across MSAs/counties occurring differentially over time correlated with housing price increases. I have not tested whether families with a higher likelihood of sending their kids to private school are sorting into locations with high housing price growth.

I do not find strong evidence to suggest wealth effects or binding credit constraints are the main driving forces behind the results. Lower-income families are expected to have more binding credit constraints. However, I find negative effects of house price changes on switching from public school to private school for low-income families. Families in the 3rd and 4th quartile of the income distribution have small and statistically insignificant effect sizes. Thus, I find no evidence to indicate strong wealth effects. My results thus point to two major driving factors behind the different estimates. First, possible improvements in school quality through increased tax revenue. Second, parents may be more willing to invest in private education at the high school level due to the heightened importance of high school performance.

3.6 Conclusion

I use housing market variation to estimate the school choice response to a change in housing wealth using individual-level data from the NLSY. I find that rising house prices increase the likelihood of parents switching their children from private to public school. This result holds across income, race, gender, and religion. My results also show increases in switching from public to private school when transitioning from middle to high school. A lack of geographic information in my data is a limitation of this study. This prevents me from accurately clustering for standard errors, separating the effects of local macroeconomic factors, analyzing local public school quality, and tracking the movement of families.

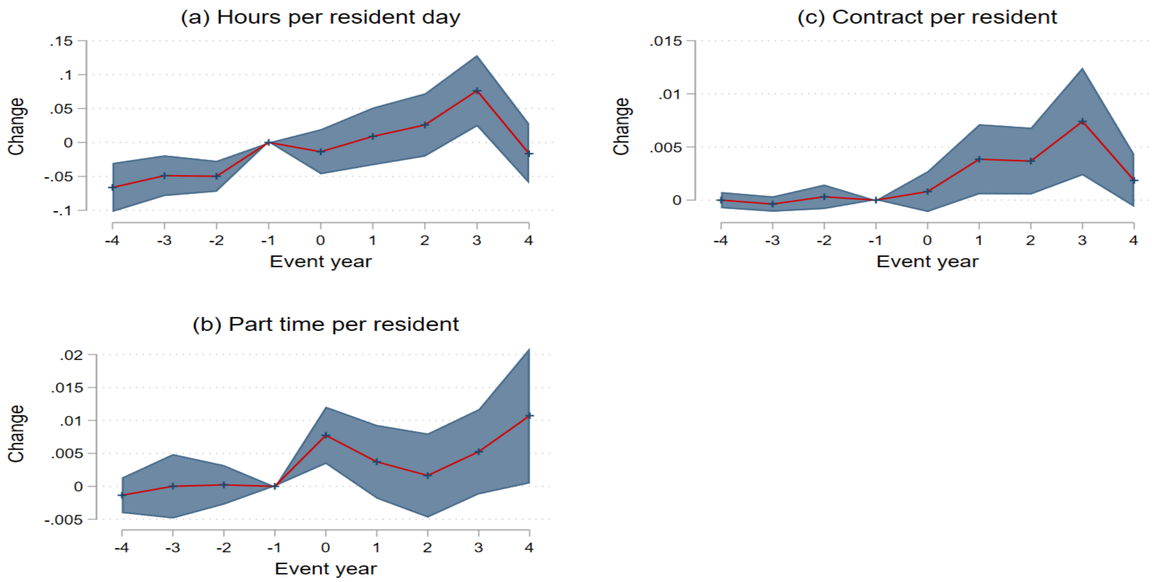
This article adds to the increasing literature examining the role of housing wealth in household behavior. Given the past volatility of the housing market and the increased focus on k-12 education in the United States, understanding how housing wealth changes affect private school enrollment is highly important. Considering the fluctuations in family resources caused by volatility in the housing market, it is likely that many families will face increasing constraints in their ability to invest in private school education in the near future. Thus, private school enrollment is sensitive to housing market fluctuations, and future research is needed on policies that can insulate school enrollment choices from variation in the housing market.

APPENDICES

APPENDIX A

CHAPTER 1 APPENDIX

Figure A.1 Staffing numbers event study

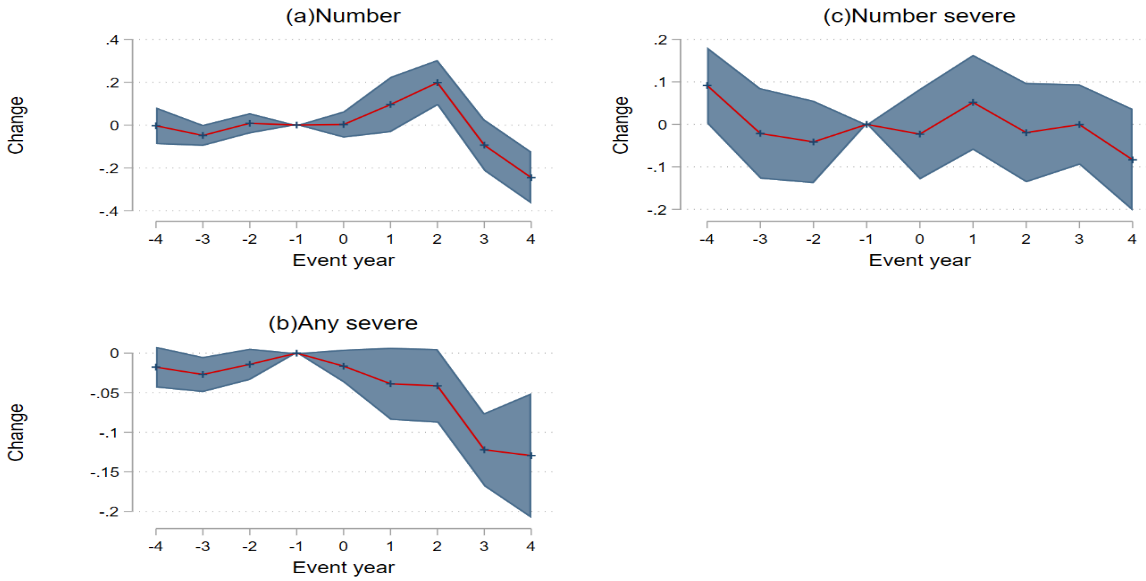


Note: Figure shows event studies with 4 prereform and 4 postreform periods. All specifications include controls for county employment rates and the elderly population share; state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels; facility and year fixed effects; Shaded areas indicate 95 percent confidence intervals with robust standard errors clustered at the county level. Data from OSCAR/CASPER 2000-2018.

Figure A.2 Health outcomes event study

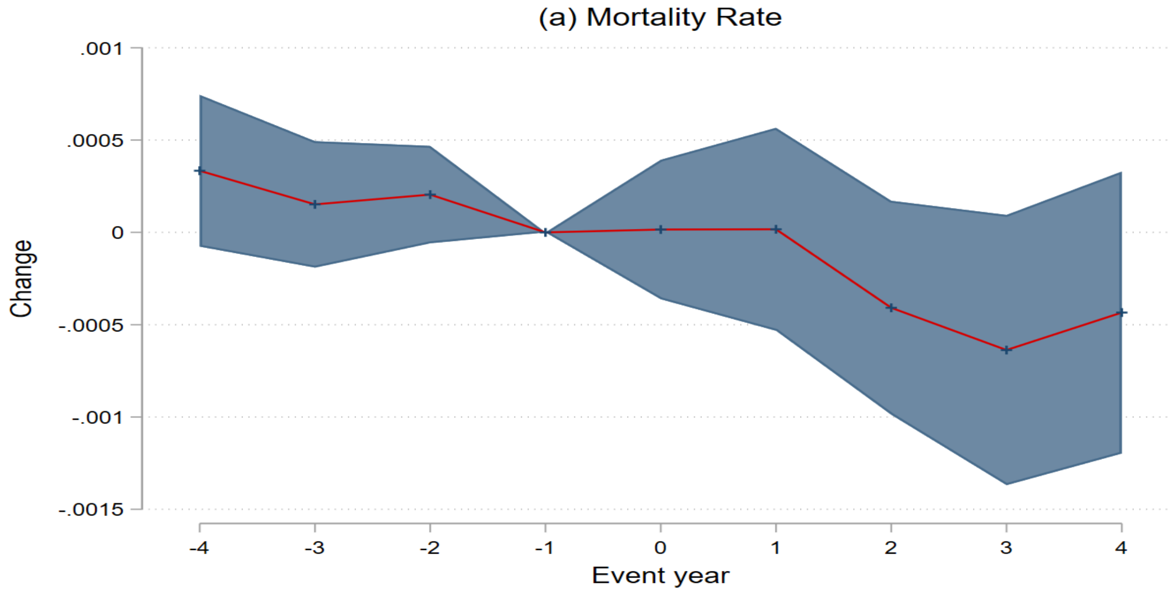
Note: Figure shows event studies with 4 prereform and 4 postreform periods. All specifications include controls for county employment rates and the elderly population share; state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels; facility and year fixed effects; Shaded areas indicate 95 percent confidence intervals with robust standard errors clustered at the county level. Data from OSCAR/CASPER 2000-2018.

Figure A.3 Violations event study



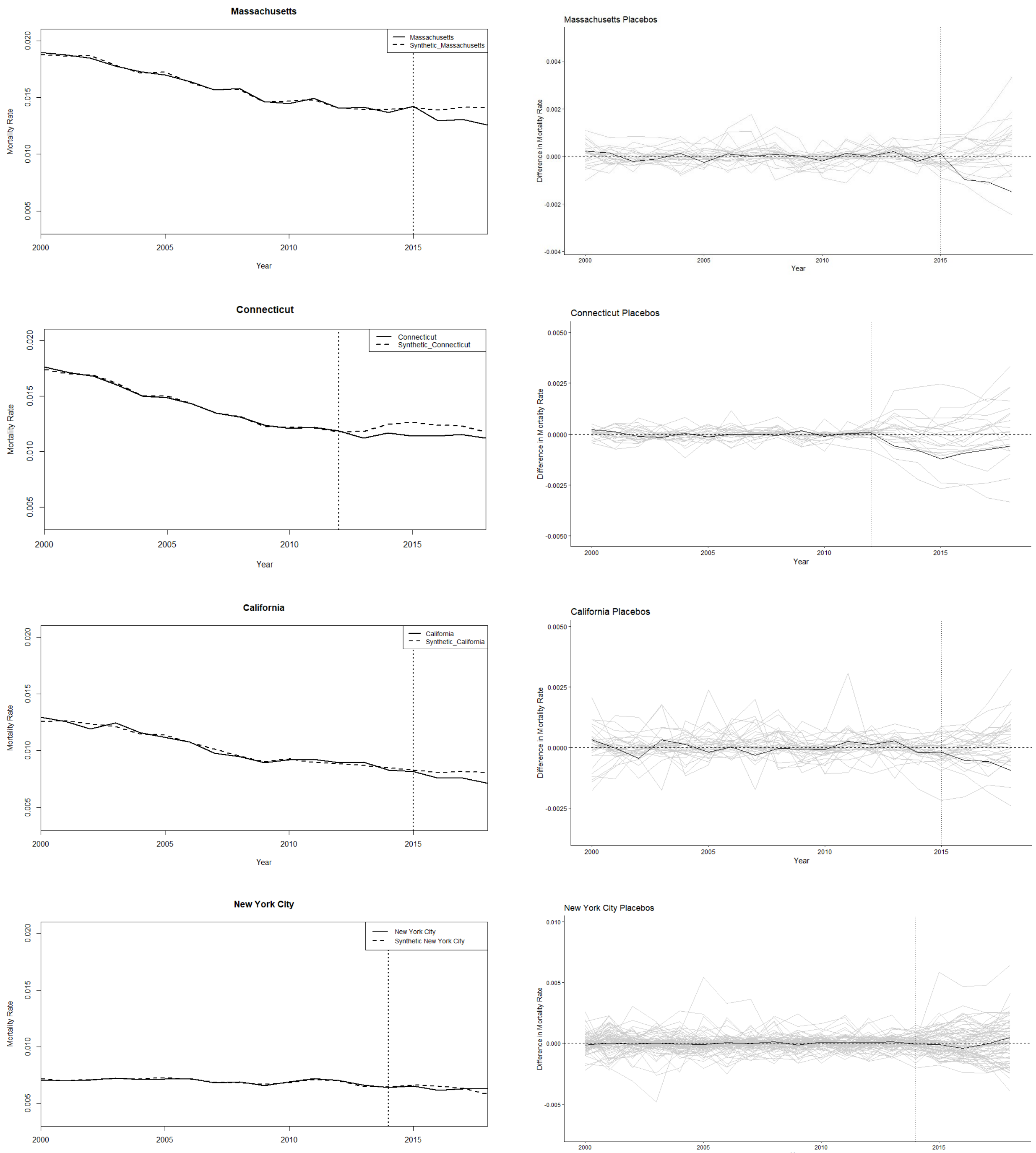
Note: Figure shows event studies with 4 prereform and 4 postreform periods. All specifications include controls for county employment rates and the elderly population share; state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels; facility and year fixed effects; Shaded areas indicate 95 percent confidence intervals with robust standard errors clustered at the county level. Data from OSCAR/CASPER 2000-2018.

Figure A.4 Nursing home elderly mortality event study



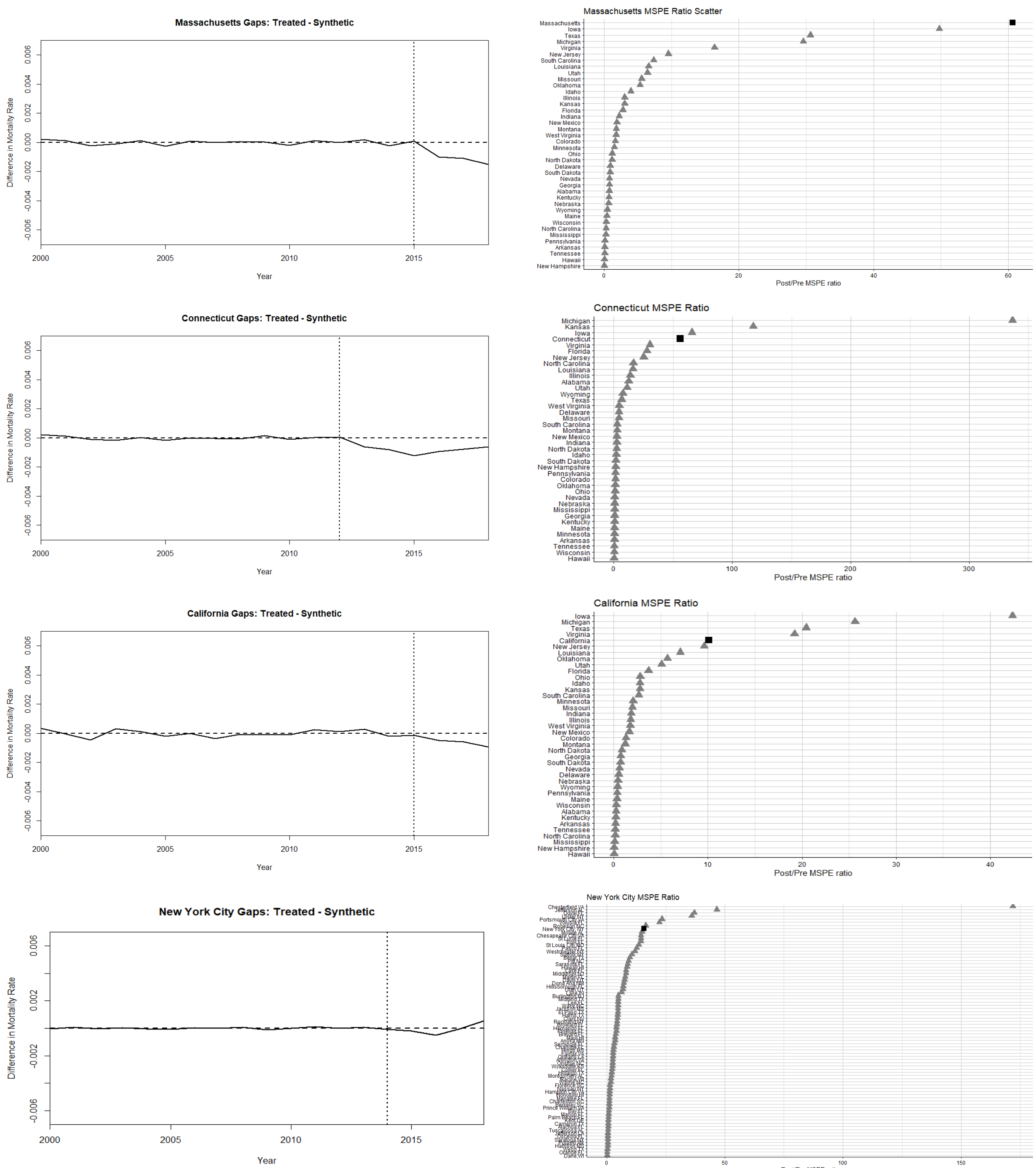
Note: Figure shows event studies with 4 prereform and 4 postreform periods. All specifications include controls for county employment rates and the elderly population share; state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels; county and year fixed effects; Shaded areas indicate 95 percent confidence intervals with robust standard errors clustered at the county level. Data from Vital Statistics microdata 2000-2018.

Figure A.5 Mortality rate in treated vs synthetic control areas



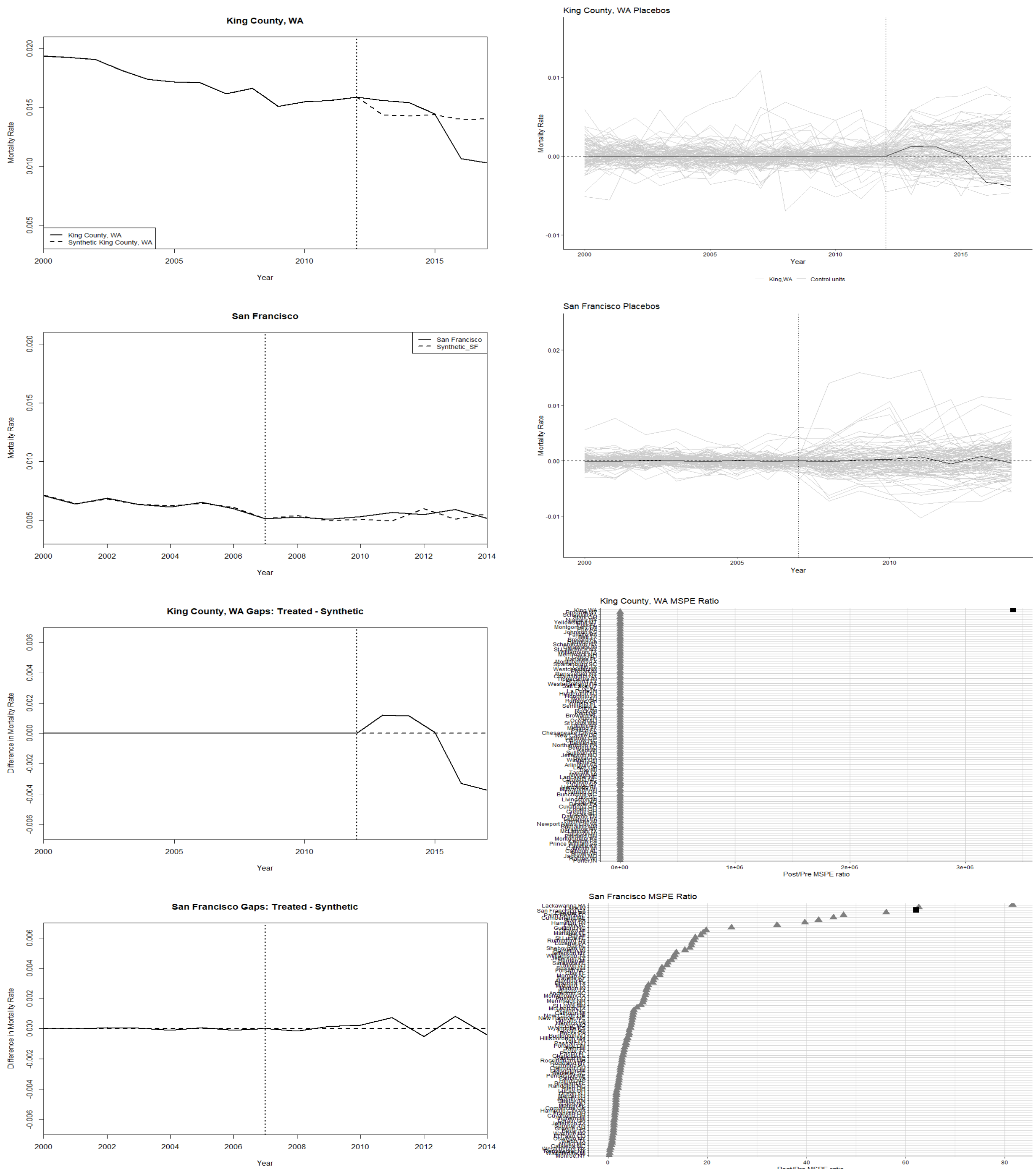
Note: The left column compares treated areas (solid lines) to the synthetic areas (dashed lines). The right column shows the difference in mortality rate between treatment and synthetic groups, and placebos for never treated areas (gray lines). The dashed vertical line indicates law enactment year. Source: Vital statistics all county mortality data.

Figure A.6 Synthetic control method - treatment effect and inference



Note: The left column shows the difference in mortality rate between the treatment and synthetic group. The dashed vertical line indicates law enactment year. Right column plots the ratio of the post-reform to pre-reform MSPE for the treatment (black square) and the never treated areas (gray triangle). Source: Vital statistics all county mortality data.

Figure A.7 Mortality rate in treated vs synthetic control areas



Note: The left column shows the difference in mortality rate between the treatment and synthetic group. The dashed vertical line indicates law enactment year. Right column plots the ratio of the post-reform to pre-reform MSPE for the treatment (black square) and the never treated areas (gray triangle). Source: Vital statistics all county mortality data.

Table A.1 Overview of PSL mandates in the U.S.

(1)	(2)	(3)	(4)
Area	Law effective	hours accrued/40 hours	Benefit under law
San Francisco	Feb 5, 2007	1.33	b/w 5 to 9 days
Connecticut	Jan 1,2012	1	up to 5 days
Seattle, WA	Sep 1,2012	1 or 1.33	b/w to 13 days
New York, NY	Apr 1, 2014	1.33	up to 40 hrs
Portland, OR	Jan 1 2014	1.33	up to 40 hrs
Jersey City, NJ	Jan 22, 2014	1.33	up to 40 hrs
Newark, NJ	May 29, 2014	1.33	b/w 24 to 40 hrs
Philadelphia, PA	May 13, 2015	1	up to 40 hrs
California	Jul 1, 2015	1.33	24 hrs minimum
Massachusetts	Jul 1, 2015	1	up to 40 hrs
Oregon	Jan 1, 2016	1.33	up to 40 hrs
Montgomery county	Oct, 2016	1.33	b/w 32 to 56 hrs
Arizona	Jul,2017	1.33	b/w 24 to 40 hrs
Maryland	Feb,2018	1.33	up to 40 hrs
Rhode Island	Jul, 2018	1.14	b/w 32 to 40 hrs
Washington	Jan 2018	1	max carryover is 40 hrs
Cook county, IL	July, 2017	1	up to 40 hrs
Chicago, IL	Jul, 2017	1	up to 40 hrs
Minneapolis	Jul, 2017	1.33	up to 48 hrs
St. Paul	Jan, 2018	1.33	up to 48 hrs
Vermont	Jan, 2017	0.77	b/w 24 and 40 hrs
New Jersey	Oct 29, 2018	1.33	up to 40 hrs

Table A.2 Nursing home and area characteristics, by treatment

	(1)	(2)
	Control Areas	Treatment Areas
VARIABLES	mean	mean
AFDC/TANF maximum	435.9	723.9
% NH residents female	70.78	67.30
Avg NH resident age	80.67	78.85
% NH residents Medicaid	60.14	61.79
Minimum Wage	7.660	9.082
State EITC Rate	0.054	0.133
Avg facility size	103.8	116.1
County HHI	0.253	0.079
Any state EITC	0.410	0.520
Share popn > 65	0.151	0.132
Share state > 65 ssi recipients	0.020	0.047
Cty unemployment	6.001	6.697

Table A.3 Nursing assistant staffing in Facilities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Full time		Part time		Contract		Hours per	
	Weekly employee hours per resident						resident day	
Paid Sick Leave	-0.017	-0.011	0.356***	0.370***	0.035	0.041	0.053***	0.058***
	(0.019)	(0.020)	(0.071)	(0.073)	(0.047)	(0.052)	(0.016)	(0.017)
Observations	225,819	225,813	225,819	225,813	225,819	225,813	245,172	245,169
Facilities	16,401		16,401		16,401		16,563	
Facility FE	YES		YES		YES		YES	
DV Mean	13.087	13.087	2.965	2.965	0.182	0.182	2.315	2.315

Note: Table shows the effect of mandated paid sick leave laws on nursing assistant staffing hours. I report results from the OSCAR/CASPER staffing reports reported by facilities to CMS, covering years 2000-2018 (columns 1-6)) and 2000-2018 (columns (7-8)). Hours per resident day is defined as the total weekly number of nursing assistant staffing hours times 35, divided by the number of residents times 7 (including direct care and administrative time). Full time employees defined as the number of nursing assistants typically working at least 35 hours a week; Part time employees defined as those typically working fewer than 35 hours a week. All specifications include controls for county employment rates and the elderly population share; and state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels. Odd numbered columns include establishment fixed effects and even numbered columns county fixed effects. Demographic controls include average resident age, facility size, and the share of residents female, and covered by Medicaid. Robust standard errors clustered by county. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table A.4 Health inspection violations

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Log(number)		Any severe		Log(number severe)	
Panel A: All health violations						
Paid Sick Leave	0.03	0.023	-0.024	-0.026*	-0.073**	-0.092***
	(0.039)	(0.038)	(0.015)	(0.015)	(0.034)	(0.026)
Observations	201,807	201,804	201,807	201,804	201,807	201,804
Facilities	16,516		16,516		16,516	
Facility FE	YES		YES		YES	
Panel B: Quality of care violations						
Paid Sick Leave	0.019	0.026	-0.011	-0.014	-0.089**	-0.102**
	(0.032)	(0.032)	(0.009)	(0.009)	(0.04)	(0.032)
Observations	201,807	201,804	201,807	201,804	201,807	201,804
Facilities	16,516		16,516		16,516	
Facility FE	YES		YES		YES	

Note: Table shows results from the state health inspection reports reported to CMS, covering years 2000-2018. Severe violations are those presenting actual harm or immediate jeopardy to residents (CMS categories G-L). Quality of care violations follow the definition in Harrington et al. (2001) to include violations in the quality of care, assessment, nursing, dietary, physician, rehabilitative services, dental, and pharmacy regulation categories. All specifications include controls for county employment rates and the elderly population share; and state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels. Odd numbered columns include facility fixed effects and even numbered columns county fixed effects. Demographic controls include average resident age, facility size, and the share of residents female, and covered by Medicaid. Robust standard errors clustered by county. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table A.5 Patient Health Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Pressure ulcers		UTI		Restraints		Psychotropic	
Paid Sick Leave	-0.615***	-0.628***	0.017	-0.004	0.847***	1.008***	-1.091***	-1.024**
	(0.124)	(0.128)	(0.134)	(0.129)	(0.150)	(0.154)	(0.422)	(0.431)
Observations	169,640	169,637	175,026	175,023	175,153	175,150	92,859	92,851
Facilities	15,872		15,896		15,902		15,246	
Facility FE	YES		YES		YES		YES	
DV Mean	5.225	5.225	6.370	6.370	2.419	2.419	19.27	19.27

Note: Table shows patient outcomes results from long-term resident assessment reports reported by facilities to CMS, covering years 2000- 2018. Reports for psychotropic medications available beginning 2005. All variables are winsorized at the 99th percentile to exclude extreme values. All specifications include controls for county employment rates and the elderly population share; and state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels. Odd numbered columns include establishment fixed effects and even numbered columns county fixed effects. Demographic controls include average resident age, facility size, and the share of residents female, and covered by Medicaid. Robust standard errors clustered by county. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table A.6 Payment methods and care needs

	(1)	(2)	(3)	(4)
Panel a:	Resident Share			
	Medicaid	Medicare	Other	Hospitalizations
Paid Sick Leave	-0.779	-0.534	1.313*	-0.00899
	(0.771)	(0.331)	(0.701)	(0.0202)
Observations	245,366	245,366	245,366	229,998
Number of establishments	16,569	16,569	16,569	16,492
DV Mean	60.60	14.83	24.57	0.967
Panel b:			Average resident care needs	
	Occupancy Rate	Successful Discharge	ADL Index	Care Index
Paid Sick Leave	0.241	-0.00513	0.320**	0.0159***
	(0.454)	(0.00316)	(0.127)	(0.00582)
Observations	245,203	93,943	245,365	245,365
Facilities	16,566	15,035	16,569	16,569
DV Mean	83.31	0.517	15.97	0.935

Note: Panel a of table shows the share of nursing home residents by payment source (columns (1) through (3)); Discharge, transfer, and occupancy rate (panel (a) column(4) and panel b column(1) and (2)) average standardized care needs (Panel (b) columns (3) and (4)); derived from resident assessment reports reported by facilities to CMS covering years 2000 through 2017, summarized in LTC focus. All specifications include controls for county employment rates and the elderly population share; and state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels; and establishment fixed effects. Demographic controls include average resident age, facility size, and the share of residents female, and covered by Medicaid. Robust standard errors clustered by county. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table A.7 Payment care needs and demographics

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Patient demographics		Patient care needs		
	% Female	% Hospital Admit	% Hypertension	% Incontinence Bladder	% Incontinence Bowel
Paid Sick Leave	-0.693***	-0.888**	-1.057	-2.354***	-0.685
	(0.139)	(0.370)	(0.876)	(0.476)	(0.479)
Observations	245,414	230,292	238,394	241,618	232,080
Facilities	16,574	16,210	16,347	16,339	16,248
DV Mean	70.13	77.37	61.61	67.55	54.57

Note: Derived from resident assessment reports reported by facilities to CMS covering years 2000 through 2017, summarized in LTC focus. All specifications include controls for county employment rates and the elderly population share; and state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels; and establishment fixed effects. Demographic controls include average resident age, facility size, and the share of residents female, and covered by Medicaid. Robust standard errors clustered by county. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table A.8 Mortality Rate

	(1)
VARIABLES	Mortality Rate
Paid Sick Leave	-0.0005**
	(0.0002043)
DV Mean	0.015

Note: Specification includes controls for county employment rates and the elderly population share; and state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels; Robust standard errors clustered by county. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table A.9 Synthetic control group method-The effect of psl mandates on mortality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\bar{Y}^1_{i,pre}$	MSPE Ratio	Rank/ #Placebos = P-value	% Treatment Effect	Level Treatment Effect	% Treatment Effect-2016	Change in Deaths -2016
Areas							
New York City	0.007	16	0.079	-0.3	0	-6.1	-414
Connecticut	0.014	56	0.1	-6.6	-0.001	-7.6	-539
California	0.01	10	0.125	-7	0	-6.5	-2522
Massachusetts	0.016	61	0.025	-8	-0.001	-7.1	-1004
Average/Sum		36	0.329	-5.5	0	-6.8	-4479
P val Irwin Hall			0				

Note: Table shows elderly mortality rate from years 2000-2018 using data Vital Statistics all county microdata. The age adjustment, defined in Equation (1.1), holds the age composition of the population fixed at its 2010 distribution; see Stevens et al. (2015). All statistics displayed here are discussed in Section 5. Column (1) displays the outcome measure in levels for each treated area averaged over all prereform years. Column (2) displays the RMSPE Ratio [RMSPE post/RMSPE pre]. Column (3) calculates the p-value of the RMSPE Ratio for all treated areas using the indicated number of placebo estimates. Columns (4) and (5) show the Percentage treatment effect and Level treatment effect. Column (6) shows change in nursing home deaths with respect to total elderly nursing home deaths in the respective area in 2014. For the first treatment area, the synthetic control group method was applied at the county level with the five boroughs of New York City as 1 county representing the treatment area of New York City. For the remaining treatment areas a synthetic control method was applied at the state level.

Table A.10 Nursing assistant staffing by provider characteristic

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Profit	Non -Profit	Multi -facility	Single	High medicaid	Low medicaid
	Panel b: Hours per resident day					
Paid Sick Leave	0.059***	0.009	0.061***	0.023	0.058***	0.033
	(0.018)	(0.025)	(0.020)	(0.019)	(0.018)	(0.021)
Observations	160,036	66,365	135,699	109,525	127,572	117,652
DV Mean	2.228	2.528	2.219	2.435	2.205	2.438

Note: I report results from the OSCAR/CASPER staffing reports reported by facilities to CMS, covering years 2000-2018 (columns 1-6)) and 2000-2018 (columns (7-8)). Hours per resident day is defined as the total weekly number of nursing assistant staffing hours times 35, divided by the number of residents times 7 (including direct care and administrative time). Full time employees defined as the number of nursing assistants typically working at least 35 hours a week; Part time employees defined as those typically working fewer than 35 hours a week. All specifications include controls for county employment rates and the elderly population share; and state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels. Odd numbered columns include facility fixed effects and even numbered columns county fixed effects. Demographic controls include average resident age, facility size, and the share of residents female, and covered by Medicaid. Robust standard errors clustered by county. *** = p < 0.01, ** = p < 0.05, * = p < 0.10.

Table A.11 Violations by provider characteristic

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Profit	Non -Profit	Multi -facility	Single	High medicaid	Low medicaid
	Panel a: Log(Number of severe violations)					
Paid Sick Leave	-0.0739*	-0.115*	-0.0923*	-0.0743*	-0.112**	-0.0115
	(0.0424)	(0.0656)	(0.0497)	(0.0440)	(0.0441)	(0.0498)
Observations	21,671	8,288	19,020	14,129	18,031	15,118
Number of Provider	9,844	3,974	7,961	6,688	8,106	7,543
	Panel b: Log(number of severe care violations)					
Paid Sick Leave	-0.0998**	-0.203***	-0.166**	-0.0367	-0.0627	-0.0483
	(0.0457)	(0.0638)	(0.0669)	(0.0510)	(0.0605)	(0.0635)
Observations	15,510	5,729	13,779	9,827	12,927	10,679

Note: Table shows results from the state health inspection reports reported to CMS, covering years 2000-2018. Severe violations are those presenting actual harm or immediate jeopardy to residents (CMS categories G-L). Quality of care violations follow the definition in Harrington et al. (2001) to include violations in the quality of care, assessment, nursing, dietary, physician, rehabilitative services, dental, and pharmacy regulation categories. All specifications include controls for county employment rates and the elderly population share; and state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels; and facility fixed effects. Demographic controls include average resident age, facility size, and the share of residents female, and covered by Medicaid. Robust standard errors clustered by county. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table A.12 Patient health outcome by provider characteristic

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Profit	Non -Profit	Multi -facility	Single	High medicaid	Low medicaid
			Panel a: Pressure ulcer			
Paid Sick Leave	-0.500***	-0.601***	-0.612***	-0.522***	-0.584***	-0.477***
	(0.134)	(0.161)	(0.148)	(0.162)	(0.148)	(0.143)
Observations	108,918	41,868	95,154	74,486	88,618	81,022
DV Mean	5.372	4.997	5.353	5.059	5.664	4.682
			Panel b: Restraints			
Paid Sick Leave	1.220***	0.441***	0.719***	1.171***	1.390***	0.478***
	(0.159)	(0.162)	(0.119)	(0.216)	(0.204)	(0.130)
Observations	112,676	43,623	97,900	77,253	91,340	83,813
DV Mean	2.468	1.544	2.321	2.547	2.693	2.083
			Panel c: Psychotropics			
Paid Sick Leave	-0.994*	-0.565	-0.659**	-1.608***	-1.559***	-0.376
	(0.526)	(0.406)	(0.331)	(0.572)	(0.534)	(0.378)
Observations	65,957	26,902	52,897	39,962	47,176	45,683
DV Mean	20.19	17.52	19.42	19.06	20.90	17.12

Note: Table shows patient outcomes results from long-term resident assessment reports reported by facilities to CMS, covering years 2000- 2018. Reports for psychotropic medications available beginning 2005. All variables are winsorized at the 99th percentile to exclude extreme values. All specifications include controls for county employment rates and the elderly population share; and state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels; and facility fixed effects. Demographic controls include average resident age, facility size, and the share of residents female, and covered by Medicaid. Robust standard errors clustered by county. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table A.13 Robustness Checks

	(1)
VARIABLES	Presence
Paid sick leave	-0.0151*
	(0.00793)
Observations	345,309
DV Mean	0.767

Note: Specification includes controls for county employment rates and the elderly population share; and state EITC parameters, the share of the elderly population receiving Supplemental Security Income, and AFDC/TANF caseloads and benefit levels; and facility fixed effects. Robust standard errors clustered by county. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table A.14 Staffing p-values adjusted for multi-collinearity

	(1)	(2)	(3)
VARIABLES	Full time	Part time	contract
Unadjusted P-value	0.164	0.001***	0.053*
Bonferroni P-value	0.164	0.004***	0.106
Sidak P-value	0.164	0.004***	0.103

*** p<0.01, ** p<0.05, * p<0.1

Table A.15 Health outcome p-values adjusted for multi-collinearity

	(1)	(2)	(3)	(4)
VARIABLES	Pressure ulcer	UTI	Restrained	Psychotropic
Unadjusted P-value	0.000***	0.899	0.000***	0.01***
Bonferroni P-value	0.000***	0.899	0.000***	0.02**
Sidak P-value	0.000***	0.877	0.000***	0.019**

*** p<0.01, ** p<0.05, * p<0.1

Table A.16 Violations p-values adjusted for multi-collinearity

	(1)	(2)	(3)	(4)	(5)	(6)
	Health violations			Care violations		
VARIABLES	#	Any severe	# severe	#	Any severe	# severe
Unadjusted P-value	0.441	0.082	0.019**	0.839	0.241	0.018**
Bonferroni P-value	0.883	0.327	0.105	0.883	0.722	0.105
Sidak P-value	0.688	0.289	0.101	0.839	0.562	0.101

*** p<0.01, ** p<0.05, * p<0.1

Table A.17 Detailed paid sick leave mandates in the U.S. effective on or before 2015

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Area	County	# beds in area in 2014	Share of beds in county	Date Law Effective	hours accrued/40 hrs	Benefit under law
San Francisco	SF	2636	100%	Feb, 2007	1.33	b/w 5 to 9 days
Seattle, WA	King	3292	53%	Sep,2012	1 or 1.33	up to 13 days
Portland	Multnomah	2526	84%	Jan, 2014	1.33	Up to 40 hrs
Jersey City, NJ	Hudson	1062	40%	Jan, 2014	1.33	up to 40 hrs
Newark, NJ	Essex	1148	24%	May, 2014	1.33	b/w 24 to 40 hrs
Philadelphia, PA	Philadelphia	7258	100%	May, 2015	1	Up to 40 hrs
New York, NY	Bronx, Kings Queens New York Richmond	44849	100%	Apr, 2014	1.33	up to 40 hrs
Connecticut		27671		Jan,2012	1	up to 5 days
Massachusetts		47517		Jul, 2015	1	up to 40 hrs
California		117781		Jul, 2015	1.33	24 hrs minimum

Table A.18 States for synthetic control group

	Connecticut	California	Massachusetts
Colorado	0.289	0	0.214
Florida	0	0.414	0
Georgia	0	0.026	0
Hawaii	0.180	0	0
Illinois	0.001	0.039	0.161
Kentucky	0	0	0.001
Maine	0.046	0	0
Minnesota	0	0	0.080
Missouri	0	0	0.116
Nebraska	0	0	0.001
Nevada	0	0.018	0
New Hampshire	0.046	0	0.086
New Jersey	0	0.288	0
North Dakota	0	0.056	0
Ohio	0	0	0.153
Pennsylvania	0	0	0.112
South Dakota	0.173	0	0
Virginia	0	0	0.002
Utah	0	0.159	0
Wisconsin	0.265	0	0.072

Note: The table shows the vector of weights that minimizes the MSPE (see Equation (1.4)) for all treated states and mortality rate as dependent variable. These weights are used to construct the synthetic control states in Figure 1 and Figure A1. The weights are also used to calculate the indicators in Table A.9. All states with positive fractions indicate the donor share employed by the SCGM to replicate the treatment county in the column header. All fractions in one column add to 1. Source: Vital Statistics all county microdata.

Table A.19 Synthetic control group method- nursing home elderly mortality rate

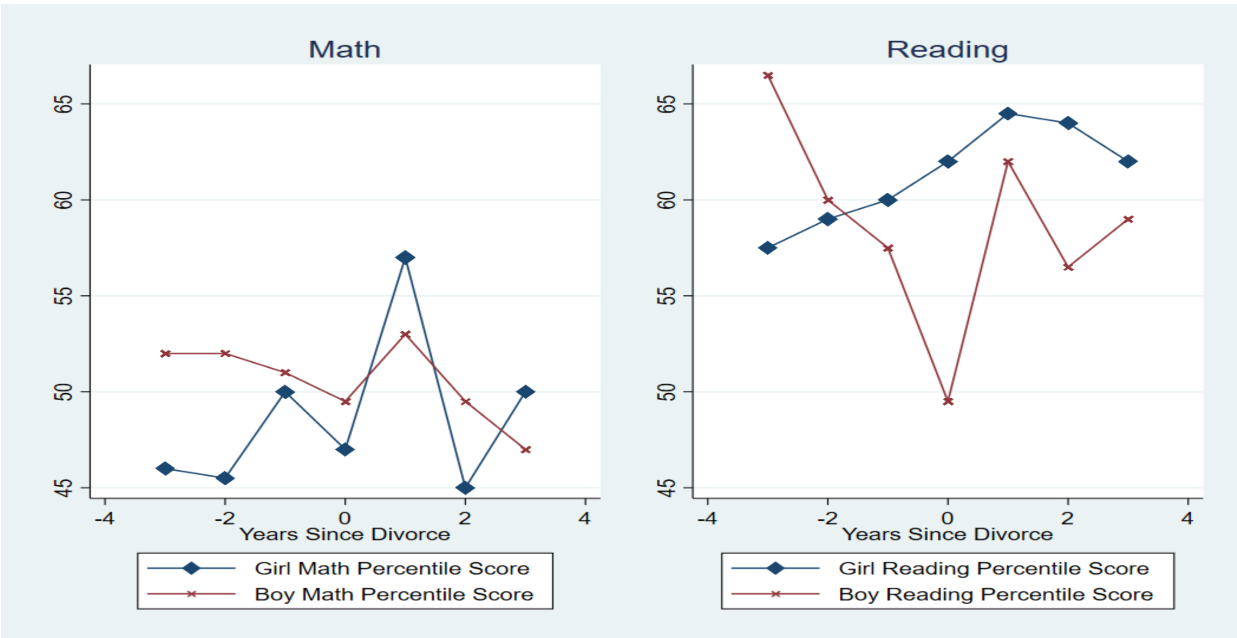
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\bar{Y}^1_{i,pre}$	MSPE Ratio	Rank-MSPE/ #Placebos = P-value	Percentage Treatment Effect	Level Treatment Effect	Percentage Treatment Effect-2016	Change in Deaths w.r.t 2016
Counties							
King County	0.017	$3 * 10^6$	0.009	-6.5	-0.001	-23.6	-604
San Francisco	0.006	62	0.03	2.4	0		
New York City	0.007	16	0.079	-0.3	0	-6.1	-414
Average/Sum		$1 * 10^6$	0.118	-1.5	0	-14.9	-1018
P val Irwin Hall			0				
States							
California	0.01	10	0.125	-7	0	-6.5	-2522
Massachusetts	0.016	61	0.025	-8	-0.001	-7.1	-1004
Connecticut	0.014	56	0.1	-6.6	-0.001	-7.6	-539
Average/Sum		42	0.25	-7.2	0	-7.1	-4065
P val Irwin Hall			0				
Total(All):							
Average/Sum		$5.6 * 10^5$	0.37	-4.4	0	-10.9	-5,083
P val Irwin Hall			0				

Note: Table shows nursing home elderly mortality rate from years 2000-2018 using data from Vital Statistics all county microdata. The age adjustment, defined in Equation 1.1, holds the age composition of the population fixed at its 2010 distribution; see Stevens et al. (2015). All statistics displayed here are discussed in Section 5. Column (1) displays the outcome measure in levels for each treated area averaged over all prereform years. Column (2) displays the MSPE Ratio [MSPE post/MSPE pre]. Column (3) calculates the p-value of the MSPE Ratio for all treated areas using the indicated number of placebo estimates. Columns (4) and (5) show the Percentage treatment effect and Level treatment effect. Column (6) shows change in nursing home deaths with respect to total elderly nursing home deaths in the area in 2014. Hudson county and Essex county are absent due to less than 50% of the beds in those two counties are part of the areas which passed the law, i.e Jersey City and Newark. Multnomah county is dropped because the state of Oregon's sick leave mandate came into effect in 2016. Philadelphia has been dropped due to there being a poor pre-reform fit. Due to California's mandate coming in to effect in 2015, San Francisco is plotted till 2014, and hence has a missing net death count for 2016.

APPENDIX B

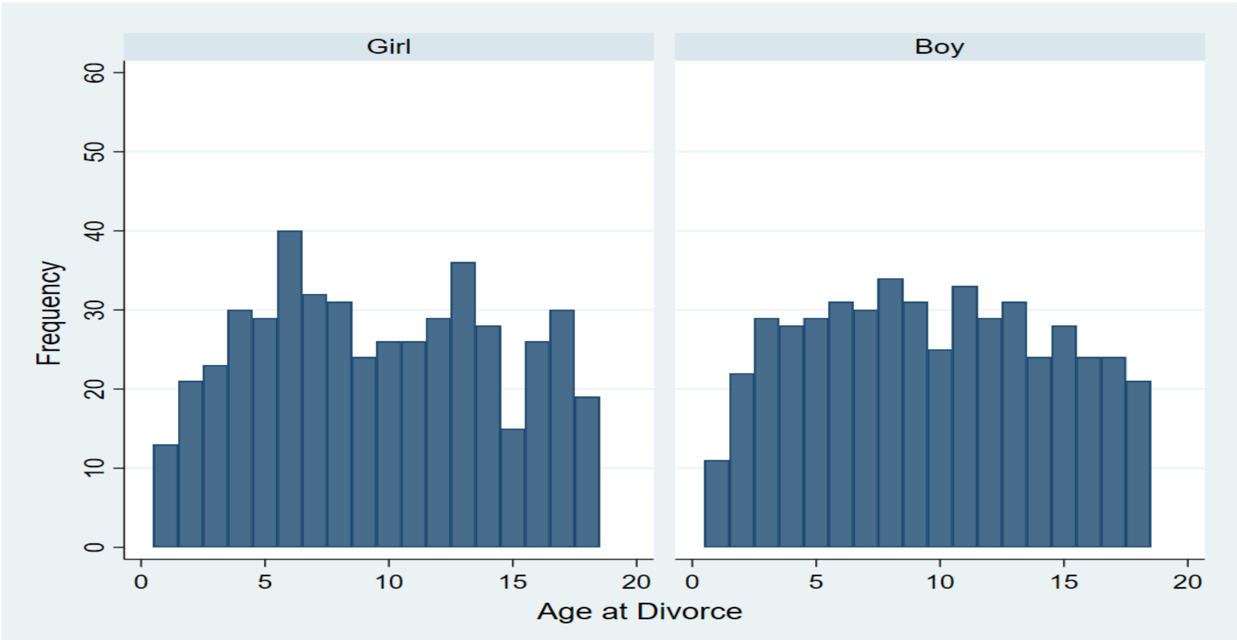
CHAPTER 2 APPENDIX

Figure B.1 Test score versus years since divorce



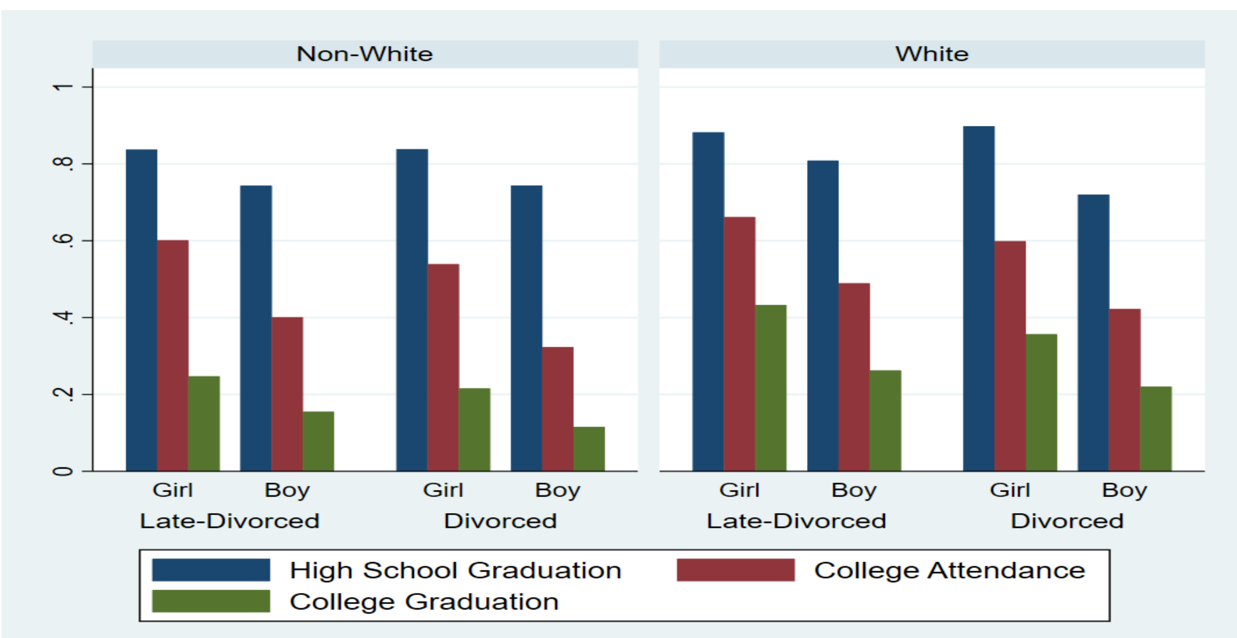
Note: The figure plots the median math and reading test scores for children of divorced families against the years since their parents have been divorced. The data used in this graph is from the CNLSY. The test scores are the Peabody Achievement scores given by the CNLSY. This figure considers divorce occurring before age 13.

Figure B.2 Distribution of ages of children at the time of parents divorce by gender



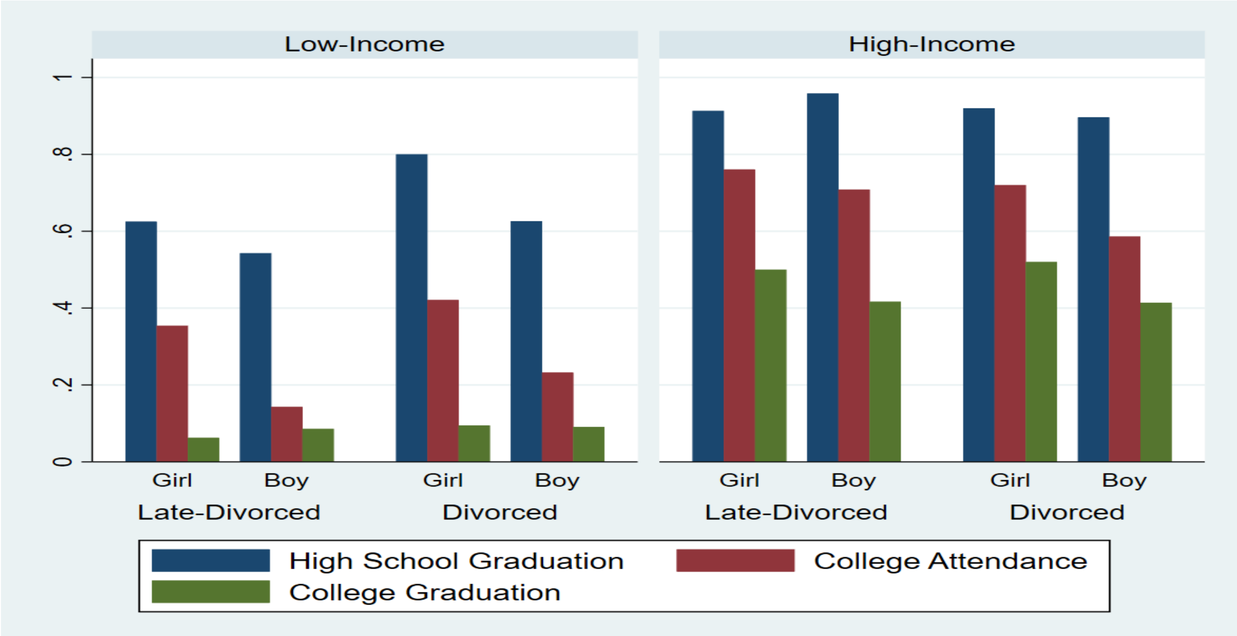
Note: This figure only considers the fixed effects sample where each family has at least two children of both genders. There are 478 girls and 484 boys whose families divorced before turning 18.

Figure B.3 Educational outcomes for each gender by race and family structure



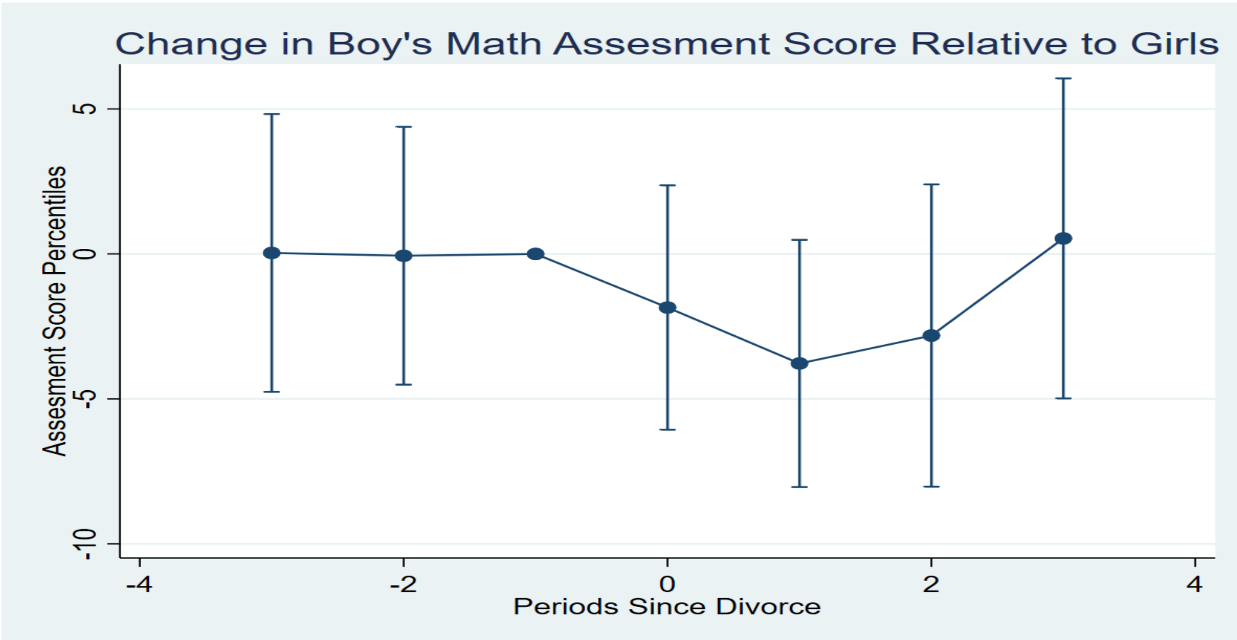
Note: Each bar gives the average educational attainment for each gender, family structure, and race category. The data comes from the CNLSY and only for children at least 19 years of age at the time of their final interview. The cutoffs for college attendance and graduation are 22 and 26, respectively.

Figure B.4 Educational outcomes for each gender by income and family structure



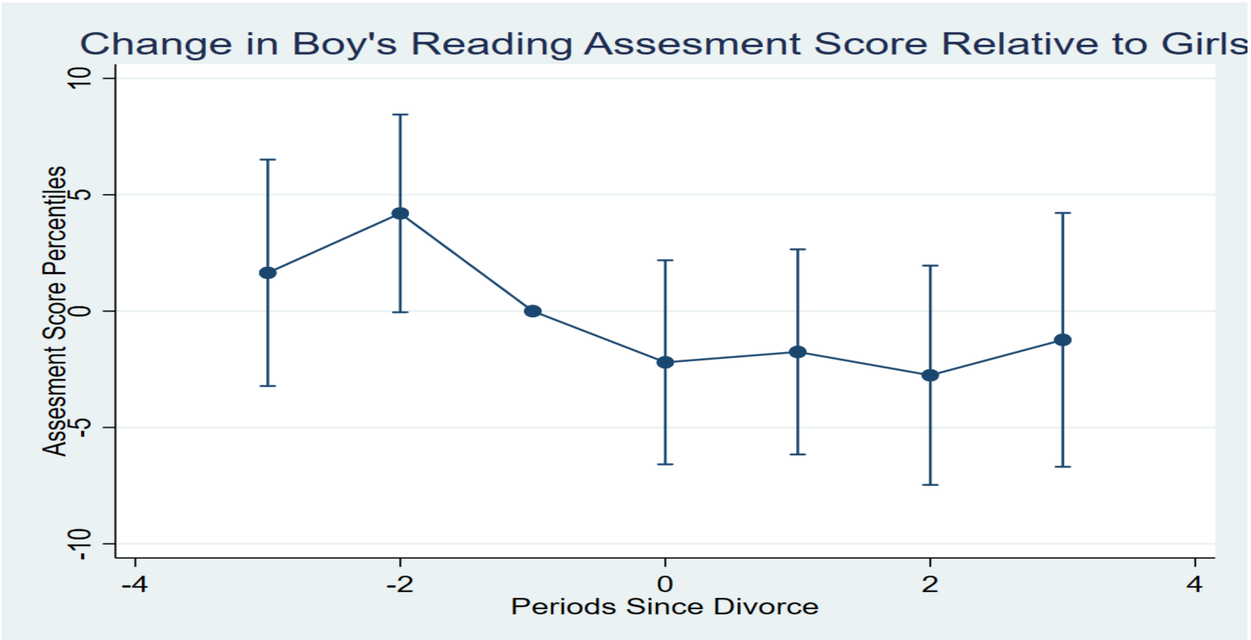
Note: Each bar gives the average educational attainment for each category of gender, family structure, and family income. High-income families are those with Permanent Income in the top quintile of the distribution and low-income families in the bottom quintile. The data comes from the CNLSY and only for children at least 19 years of age at the time of their final interview. The cutoffs for college attendance and graduation are 22 and 26, respectively.

Figure B.5 Event study for math test scores



Note: This figure reports estimates of event study regressions, including indicator variables for every two years before and after divorce, and another set of indicators interacted with a dummy variable for gender. The event study regression specification includes sibling fixed effects. The blue vertical bars represent confidence intervals.

Figure B.6 Event study for reading test scores



Note: This figure reports estimates of event study regressions. These regressions include indicator variables for two-year periods before and after divorce and another set of indicators interacted with a dummy variable for gender. The event study regression specification includes sibling fixed effects. The blue vertical bars represent confidence intervals.

Table B.1 Summary stats by gender & family structure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	EarlyDivorced			LateDivorced			NeverDivorced		
	Girl	Boy		Girl	Boy		Girl	Boy	
	Mean	Mean	Raw-Diff	Mean	Mean	Raw-Diff	Mean	Mean	Raw-Diff
HSGrad	0.868	0.734	0.134	0.856	0.771	0.085	0.910	0.842	0.068
CollegeAtt	0.563	0.373	0.190	0.626	0.439	0.187	0.661	0.583	0.078
CollegeGrad	0.286	0.171	0.115	0.325	0.201	0.124	0.400	0.300	0.100
Idle	0.145	0.122	0.023	0.170	0.110	0.060	0.163	0.120	0.043
Crime	0.245	0.443	-0.198	0.187	0.412	-0.225	0.130	0.299	-0.169
PoorHealth	0.157	0.107	0.050	0.161	0.131	0.030	0.0917	0.0604	0.0313

Note: Crime is defined as one if the respondent was ever convicted, been on probation, sentenced, or been in jail. A respondent is classified as Idle if they are not enrolled in an educational institution or earning positive wages in the last wave they show up in. A respondent is classified as having poor health if, on average, they report their health as less than good.

Table B.2 Educational outcomes regression values

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HSGrad		CollegeAtt		CollegeGrad	
EarlyDivorce	-0.004	-0.014	-0.086**	-0.091	-0.015	-0.005
	[0.027]	[0.039]	[0.040]	[0.059]	[0.046]	[0.062]
Boy	-0.095***	-0.092***	-0.212***	-0.159***	-0.128***	-0.110***
	[0.032]	[0.020]	[0.043]	[0.027]	[0.044]	[0.028]
BoyxEarlyDivorce	-0.046	-0.064**	-0.015	-0.080*	0.011	-0.011
	[0.042]	[0.032]	[0.057]	[0.043]	[0.061]	[0.047]
Premature	0.005	-0.019	-0.003	-0.009	0.015	-0.003
	[0.016]	[0.022]	[0.025]	[0.031]	[0.029]	[0.037]
Birth Order	-0.069***	-0.017	-0.097***	-0.044**	-0.088***	-0.051**
	[0.007]	[0.017]	[0.009]	[0.021]	[0.012]	[0.025]
BoyxWhite	0.006	0.021	0.051	0.068**	-0.054	-0.024
	[0.026]	[0.023]	[0.039]	[0.034]	[0.047]	[0.041]
NeverDivorce	0.009		-0.031		0.089**	
	[0.022]		[0.034]		[0.040]	
BoyxNeverDivorce	0.019		0.098**		0.053	
	[0.034]		[0.048]		[0.054]	
Observations	2,842	2,842	2,351	2,351	1,466	1,466
Sibling FE	NO	YES	NO	YES	NO	YES

Note:Robust Standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1 The regression with college graduation as outcome only considers respondents who were 26 and over at the time of their last interview. Similarly, the cutoff is 22 for college attendance. All races are that of the child. In every interview, the CNLSY asks its respondents their highest educational qualification. Some correspondents mention that they are still in college, and their highest qualification to date is college attendance. A respondent is classified as premature if he/she was born early or late.

Table B.3 Non-educational outcomes regression values

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Crime		Idle		PoorHealth	
EarlyDivorce	0.078**	0.032	-0.017	-0.029	0.001	0.044
	[0.032]	[0.047]	[0.029]	[0.041]	[0.029]	[0.040]
Boy	0.256***	0.227***	-0.048	-0.023	-0.045	-0.036*
	[0.038]	[0.024]	[0.030]	[0.021]	[0.030]	[0.018]
BoyxEarlyDivorce	-0.034	0.013	0.030	0.024	-0.015	-0.030
	[0.050]	[0.039]	[0.038]	[0.031]	[0.039]	[0.031]
Premature	0.014	-0.059**	0.038**	0.029	0.011	0.014
	[0.020]	[0.030]	[0.017]	[0.022]	[0.015]	[0.021]
BirthOrder	0.036***	-0.012	0.039***	0.026	0.021***	0.010
	[0.008]	[0.019]	[0.007]	[0.018]	[0.006]	[0.015]
BoyxWhite	-0.051	-0.084***	-0.020	-0.032	0.024	0.022
	[0.032]	[0.031]	[0.026]	[0.026]	[0.023]	[0.023]
	[0.002]	[0.006]	[0.002]	[0.006]	[0.001]	[0.005]
Never Divorce	-0.016		0.009		-0.062***	
	[0.026]		[0.025]		[0.024]	
BoyxNeverDivorce	-0.056		0.019		0.004	
	[0.040]		[0.032]		[0.031]	
Observations	2,841	2,841	2,841	2,841	2,841	2,841
Sibling FE	NO	YES	NO	YES	NO	YES

Note: Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. A respondent is classified as premature if he/she was born early or late. The race is that of the child.

Table B.4 Educational outcomes regression values, by race

	(1)	(2)	(3)	(4)	(5)	(6)
	White	Non-White	White	NonWhite	White	NonWhite
VARIABLES	HSGrad		CollegeAtt		CollegeGrad	
EarlyDivorce	-0.047	-0.013	-0.028	-0.166**	-0.033	0.034
	[0.055]	[0.055]	[0.080]	[0.084]	[0.108]	[0.074]
Boy	-0.062***	-0.102***	-0.095***	-0.158***	-0.144***	-0.097***
	[0.014]	[0.021]	[0.027]	[0.028]	[0.039]	[0.029]
BoyxEarlyDivorce	-0.106**	-0.022	-0.073	-0.087	0.039	-0.066
	[0.042]	[0.049]	[0.059]	[0.062]	[0.077]	[0.059]
Birth Order	-0.021	-0.011	-0.068**	-0.027	-0.062	-0.038
	[0.023]	[0.023]	[0.033]	[0.028]	[0.047]	[0.029]
Premature	0.007	-0.049	-0.034	0.006	0.030	-0.019
	[0.024]	[0.036]	[0.045]	[0.043]	[0.065]	[0.044]
Observations	1,397	1,445	1,130	1,221	648	818
Sibling FE	YES	YES	YES	YES	YES	YES

Note: Robust Standard errors in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. A respondent is classified as premature if he/she was born early or late. The race is that of the child. The regression with college graduation as outcome only considers respondents who were 26 and over at the time of their last interview. Similarly, the cutoff is 22 for college attendance.

Table B.5 Educational outcomes regression values, by income

	(1)	(2)	(3)	(4)	(5)	(6)
	Hi-Income	Lo-Income	Hi-Income	Lo-Income	Hi-Income	Lo-Income
VARIABLES	HSGrad		CollegeAtt		CollegeGrad	
EarlyDivorce	-0.004	0.136	0.054	0.047	-0.301	0.070
	[0.039]	[0.121]	[0.064]	[0.133]	[0.221]	[0.118]
Boy	0.005	-0.128**	0.014	-0.203***	-0.004	-0.103*
	[0.028]	[0.052]	[0.075]	[0.060]	[0.142]	[0.058]
BoyxEarlyDivorce	0.001	-0.097	-0.027	-0.107	0.407	0.014
	[0.067]	[0.093]	[0.123]	[0.103]	[0.257]	[0.076]
Birth Order	-0.023	-0.067	-0.013	-0.044	-0.123	0.028
	[0.030]	[0.045]	[0.052]	[0.042]	[0.089]	[0.053]
BoyxWhite	-0.006	0.023	-0.023	-0.054	-0.124	0.096
	[0.032]	[0.095]	[0.083]	[0.109]	[0.151]	[0.112]
Premature	-0.006	-0.015	-0.021	-0.049	0.010	-0.126*
	[0.022]	[0.078]	[0.054]	[0.077]	[0.119]	[0.069]
	[0.012]	[0.017]	[0.020]	[0.018]	[0.034]	[0.032]
Observations	552	591	412	486	223	282
Sibling FE	YES	YES	YES	YES	YES	YES

Note: Robust Standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. A respondent is classified as premature if he/she was born early or late. The race is that of the child. The regression with college graduation as outcome only considers respondents who were 26 and over at the time of their last interview. Similarly, the cutoff is 22 for college attendance.

Table B.6 Exposure time linear regression

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HSGrad		CollegeAtt		CollegeGrad	
Exposure	-0.001	0.003	-0.006*	-0.004	-0.003	0.003
	[0.003]	[0.004]	[0.004]	[0.006]	[0.004]	[0.006]
Boy	-0.116***	-0.095***	-0.208***	-0.159***	-0.142***	-0.113***
	[0.034]	[0.020]	[0.045]	[0.027]	[0.047]	[0.029]
BoyxExposure	-0.000	-0.004	-0.002	-0.006*	0.003	0.001
	[0.004]	[0.003]	[0.005]	[0.004]	[0.006]	[0.004]
Never Divorce	0.004		-0.026		0.079*	
	[0.023]		[0.036]		[0.042]	
BoyxNeverDivorce	0.041		0.093*		0.068	
	[0.035]		[0.049]		[0.057]	
Premature	0.006	-0.019	-0.003	-0.008	0.015	-0.002
	[0.016]	[0.022]	[0.025]	[0.031]	[0.029]	[0.037]
Birth Order	-0.069***	-0.017	-0.096***	-0.042**	-0.088***	-0.050**
	[0.007]	[0.017]	[0.009]	[0.021]	[0.012]	[0.025]
BoyxWhite	0.005	0.021	0.052	0.067**	-0.055	-0.026
	[0.026]	[0.023]	[0.039]	[0.034]	[0.047]	[0.041]
Observations	2,842	2,842	2,351	2,351	1,466	1,466
Sibling FE	No	YES	No	YES	No	YES

Note: Robust Standard errors in brackets.*** p<0.01, ** p<0.05, * p<0.1. Exposure is a linear function of 18-age at divorce. The regression with college graduation as outcome only considers respondents who were 26 and over at the time of their last interview. Similarly, the cutoff is 22 for college attendance. A respondent is coded as premature if he/she is born early or late.

Table B.7 Age group exposure time model

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HSGrad		CollegeAtt		CollegeGrad	
Young	-0.030	0.018	-0.091	-0.065	-0.031	0.015
	[0.041]	[0.059]	[0.058]	[0.083]	[0.065]	[0.092]
Mid Child	0.006	-0.023	-0.060	-0.024	-0.011	-0.012
	[0.032]	[0.053]	[0.048]	[0.073]	[0.056]	[0.085]
Teenager	0.004	0.033	0.024	0.138**	-0.042	-0.019
	[0.033]	[0.047]	[0.048]	[0.063]	[0.053]	[0.073]
Boy	-0.073*	-0.085***	-0.174***	-0.141***	-0.104*	-0.100***
	[0.039]	[0.021]	[0.052]	[0.028]	[0.056]	[0.030]
NeverDivorce	0.009		-0.018		0.074	
	[0.027]		[0.041]		[0.048]	
BoyxNeverDivorce	-0.002		0.060		0.030	
	[0.040]		[0.056]		[0.064]	
BoyxYoung	0.021	-0.031	-0.060	-0.069	0.077	0.092
	[0.060]	[0.059]	[0.079]	[0.075]	[0.088]	[0.079]
BoyxMidchild	-0.092*	-0.074*	-0.036	-0.075	-0.054	-0.086
	[0.050]	[0.040]	[0.065]	[0.053]	[0.073]	[0.056]
BoyxTeenager	-0.059	-0.055	-0.083	-0.122**	-0.055	-0.040
	[0.051]	[0.040]	[0.067]	[0.053]	[0.069]	[0.053]
Observations	2,842	2,842	2,351	2,351	1,466	1,466
Sibling FE	No	YES	No	YES	No	YES

Note: Robust Standard errors in brackets.*** p<0.01, ** p<0.05, * p<0.1. The three age groups, young, mid-child, and teen, are for children in age groups (0-4),(5-12), and (13-18), respectively. Standard controls are included in the model but not provided in the table. The regression with college graduation as outcome only considers respondents who were 26 and over at the time of their last interview. Similarly, the cutoff is 22 for college attendance.

Table B.8 Placebo tests using assessment scores

	(1)	(2)	(3)	(4)
VARIABLES	BPI-5	HOME-5	PIATMT-5	PIATRR-5
Boy	3.647**	-2.594*	-1.025	-5.777***
	[1.636]	[1.354]	[1.802]	[1.663]
BoyxEarlyDivorce	0.936	1.643	-2.692	2.589
	[3.729]	[3.087]	[3.992]	[3.562]
EarlyDivorce	0.783	-3.053	1.804	-7.068*
	[3.646]	[3.457]	[4.319]	[4.082]
Boyxwhite	-0.952	0.313	0.547	1.389
	[2.029]	[1.771]	[2.242]	[2.082]
BirthOrder	-1.119	-0.583	-1.701	-2.744**
	[1.128]	[0.915]	[1.173]	[1.193]
Premature	0.020	-1.091	-1.026	-2.878
	[1.721]	[1.561]	[1.889]	[1.800]
Observations	2,648	2,663	2,585	2,542
R-squared	0.706	0.791	0.667	0.675
Sibling FE	YES	YES	YES	YES

Note: Robust Standard errors in brackets.*** $p<0.01$, ** $p<0.05$, * $p<0.1$. PIATMT-5 is the Peabody Math achievement score for a child at ages 5 or 6. It is a percentile score. A child is classified as premature if he/she has been born early or late. The race is that of the child.

Table B.9 Standardized educational outcome regression values

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HSGrad		CollegeAtt		CollegeGrad	
EarlyDivorce	-0.011	-0.038	-0.174**	-0.184	-0.033	-0.010
	[0.074]	[0.109]	[0.081]	[0.119]	[0.100]	[0.134]
Boy	-0.261***	-0.254***	-0.427***	-0.320***	-0.274***	-0.236***
	[0.088]	[0.054]	[0.087]	[0.054]	[0.095]	[0.061]
BoyxEarlyDivorce	-0.128	-0.178**	-0.031	-0.161*	0.023	-0.024
	[0.117]	[0.089]	[0.115]	[0.087]	[0.132]	[0.101]
Premature	0.014	-0.054	-0.005	-0.017	0.032	-0.007
	[0.045]	[0.060]	[0.050]	[0.063]	[0.063]	[0.079]
Birth Order	-0.190***	-0.048	-0.195***	-0.088**	-0.190***	-0.109**
	[0.020]	[0.047]	[0.019]	[0.042]	[0.025]	[0.054]
BoyxWhite	0.018	0.059	0.103	0.137**	-0.115	-0.052
	[0.072]	[0.065]	[0.078]	[0.068]	[0.102]	[0.088]
Never Divorce	0.026		-0.062		0.191**	
	[0.061]		[0.069]		[0.087]	
BoyxNeverDivorce	0.053		0.197**		0.114	
	[0.093]		[0.096]		[0.117]	
Observations	2,842	2,842	2,351	2,351	1,466	1,466
Sibling FE	NO	YES	NO	YES	NO	YES

Note: Robust Standard errors in brackets.*** p<0.01, ** p<0.05, * p<0.1. The regression with college graduation as outcome only considers respondents who were 26 and over at the time of their last interview. Similarly the cutoff is 22 for college attendance. All races are that of the child. In every interview the CNLSY asks its respondents their highest educational qualification. There are correspondents who mention that they are still in college and their highest qualification till date is college attendance. A respondent is classified as premature if he/she was born early or late.

Table B.10 Placebo tests with standardized assessment scores

	(1)	(2)	(3)	(4)
VARIABLES	BPI-5	HOME-5	PIATMT-5	PIATRR-5
Boy	0.129**	-0.090*	-0.037	-0.221***
	[0.058]	[0.047]	[0.064]	[0.064]
BoyxEarlyDivorce	0.033	0.057	-0.096	0.099
	[0.132]	[0.108]	[0.143]	[0.136]
EarlyDivorce	0.028	-0.106	0.065	-0.270*
	[0.129]	[0.121]	[0.154]	[0.156]
BoyxWhite	-0.034	0.011	0.020	0.053
	[0.072]	[0.062]	[0.080]	[0.080]
BirthOrder	-0.040	-0.020	-0.061	-0.105**
	[0.040]	[0.032]	[0.042]	[0.046]
Premature	0.001	-0.038	-0.037	-0.110
	[0.061]	[0.054]	[0.068]	[0.069]
Observations	2,648	2,663	2,585	2,542
Sibling FE	YES	YES	YES	YES

Note: Robust Standard errors in brackets.*** $p<0.01$, ** $p<0.05$, * $p<0.1$. I am taking two year gaps because the surveys are conducted every two years. So a peabody math test score at age 7 essentially means ages 7 & 8 as a child will only be tested in alternate years. A child is classified as premature if he/she has been born early or late. The race is that of the child.

Table B.11 Educational outcome regression values for the non-incarcerated sample

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HSGrad		CollegeAtt		CollegeGrad	
EarlyDivorce	-0.007	-0.014	-0.088**	-0.092	-0.016	-0.005
	[0.027]	[0.039]	[0.041]	[0.059]	[0.047]	[0.062]
Boy	-0.098***	-0.093***	-0.209***	-0.155***	-0.129***	-0.110***
	[0.032]	[0.020]	[0.043]	[0.027]	[0.044]	[0.028]
BoyxEarlyDivorce	-0.043	-0.064**	-0.014	-0.077*	0.011	-0.011
	[0.042]	[0.032]	[0.057]	[0.043]	[0.062]	[0.048]
Premature	0.005	-0.019	-0.003	-0.008	0.013	-0.002
	[0.017]	[0.022]	[0.025]	[0.031]	[0.029]	[0.037]
Birth Order	-0.069***	-0.018	-0.097***	-0.046**	-0.087***	-0.053**
	[0.007]	[0.017]	[0.009]	[0.021]	[0.012]	[0.025]
BoyxWhite	0.006	0.021	0.048	0.062*	-0.053	-0.027
	[0.026]	[0.023]	[0.039]	[0.034]	[0.047]	[0.041]
NeverDivorce	0.007		-0.032		0.084**	
	[0.022]		[0.034]		[0.041]	
BoyxNeverDivorce	0.023		0.096**		0.053	
	[0.034]		[0.048]		[0.055]	
White	0.036**		0.059**		0.171***	
	[0.017]		[0.028]		[0.036]	
Observations	2,831	2,831	2,341	2,341	1,456	1,456
Sibling FE	NO	YES	NO	YES	NO	YES

Note: Robust Standard errors in brackets.*** p<0.01, ** p<0.05, * p<0.1. The regression with college graduation as outcome only considers respondents who were 26 and over at the time of their last interview. Similarly the cutoff is 22 for college attendance. All races are that of the child. In every interview the CNLSY asks its respondents their highest educational qualification. There are correspondents who mention that they are still in college and their highest qualification till date is college attendance. A respondent is classified as premature if he/she was born early or late.

Table B.12 Educational outcome regression values for the non-attrited sample

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HSGrad		CollegeAtt		CollegeGrad	
EarlyDivorce	-0.004	-0.014	-0.086**	-0.092	-0.015	-0.005
	[0.027]	[0.039]	[0.040]	[0.059]	[0.046]	[0.062]
Boy	-0.094***	-0.092***	-0.212***	-0.159***	-0.128***	-0.110***
	[0.032]	[0.020]	[0.043]	[0.027]	[0.044]	[0.028]
BoyxEarlyDivorce	-0.047	-0.065**	-0.017	-0.080*	0.011	-0.011
	[0.042]	[0.032]	[0.057]	[0.043]	[0.061]	[0.047]
Premature	0.005	-0.020	-0.003	-0.009	0.015	-0.003
	[0.017]	[0.022]	[0.025]	[0.031]	[0.029]	[0.037]
BirthOrder	-0.069***	-0.017	-0.096***	-0.044**	-0.088***	-0.051**
	[0.007]	[0.017]	[0.009]	[0.021]	[0.012]	[0.025]
BoyxWhite	0.006	0.021	0.050	0.068**	-0.054	-0.024
	[0.026]	[0.023]	[0.039]	[0.034]	[0.047]	[0.041]
NeverDivorce	0.009		-0.031		0.089**	
	[0.022]		[0.034]		[0.040]	
BoyxNeverDivorce	0.019		0.098**		0.053	
	[0.034]		[0.048]		[0.054]	
White	0.034**		0.057**		0.173***	
	[0.017]		[0.028]		[0.036]	
Observations	2,841	2,841	2,350	2,350	1,466	1,466
Sibling FE	NO	YES	NO	YES	NO	YES

Note: Robust Standard errors in brackets.*** p<0.01, ** p<0.05, * p<0.1. The regression with college graduation as outcome only considers respondents who were 26 and over at the time of their last interview. Similarly the cutoff is 22 for college attendance. All races are that of the child. In every interview the CNLSY asks its respondents their highest educational qualification. There are correspondents who mention that they are still in college and their highest qualification till date is college attendance. A respondent is classified as premature if he/she was born early or late.

Table B.13 Educational outcome regression values using alternate cutoffs, 1

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HSGrad		CollegeAtt		CollegeGrad	
EarlyDivorce	-0.032	0.097*	-0.051	0.072	-0.087	-0.012
	[0.044]	[0.058]	[0.062]	[0.072]	[0.069]	[0.079]
Boy	-0.128***	-0.104***	-0.212***	-0.167***	-0.140***	-0.118***
	[0.026]	[0.019]	[0.036]	[0.026]	[0.036]	[0.026]
BoyxEarlyDivorce	0.085	-0.060	-0.033	-0.140*	0.140	0.092
	[0.064]	[0.072]	[0.087]	[0.083]	[0.091]	[0.092]
Premature	0.007	-0.017	-0.004	-0.007	0.017	-0.000
	[0.017]	[0.022]	[0.025]	[0.031]	[0.029]	[0.037]
Birth Order	-0.070***	-0.018	-0.097***	-0.042**	-0.088***	-0.052**
	[0.007]	[0.017]	[0.009]	[0.021]	[0.012]	[0.025]
BoyxWhite	0.004	0.022	0.049	0.066*	-0.058	-0.027
	[0.026]	[0.023]	[0.039]	[0.034]	[0.047]	[0.041]
NeverDivorce	0.007		0.007		0.084**	
	[0.018]		[0.029]		[0.036]	
BoyxNeverDivorce	0.053*		0.099**		0.068	
	[0.028]		[0.041]		[0.048]	
Observations	2,842	2,842	2,351	2,351	1,466	1,466
Sibling FE	NO	YES	NO	YES	NO	YES

Note: Robust Standard errors in brackets.*** p<0.01, ** p<0.05, * p<0.1. The regression with college graduation as outcome only considers respondents who were 26 and over at the time of their last interview. Similarly the cutoff is 22 for college attendance. All races are that of the child. In every interview the CNLSY asks its respondents their highest educational qualification. There are correspondents who mention that they are still in college and their highest qualification till date is college attendance. A respondent is classified as premature if he/she was born early or late.

Table B.14 Educational outcome regression values using alternate cutoffs, 2

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HSGrad		CollegeAtt		CollegeGrad	
EarlyDivorce	0.009	0.020	-0.053	0.002	-0.039	-0.015
	[0.028]	[0.036]	[0.041]	[0.052]	[0.046]	[0.060]
Boy	-0.094***	-0.092***	-0.222***	-0.162***	-0.139***	-0.117***
	[0.035]	[0.020]	[0.047]	[0.028]	[0.049]	[0.029]
BoyxEarlyDivorce	-0.039	-0.051*	0.005	-0.051	0.028	0.016
	[0.043]	[0.030]	[0.058]	[0.041]	[0.061]	[0.045]
Premature	0.005	-0.018	-0.003	-0.007	0.015	-0.003
	[0.017]	[0.022]	[0.025]	[0.031]	[0.029]	[0.037]
Birth Order	-0.069***	-0.017	-0.097***	-0.042**	-0.088***	-0.051**
	[0.007]	[0.017]	[0.009]	[0.021]	[0.011]	[0.025]
BoyxWhite	0.006	0.020	0.050	0.066*	-0.053	-0.025
	[0.026]	[0.023]	[0.039]	[0.034]	[0.047]	[0.041]
Never Divorce	0.017		-0.019		0.074*	
	[0.025]		[0.037]		[0.043]	
Boy Never Divorce	0.019		0.108**		0.065	
	[0.037]		[0.051]		[0.058]	
Observations	2,842	2,842	2,351	2,351	1,466	1,466
Sibling FE	NO	YES	NO	YES	NO	YES

Note: Robust Standard errors in brackets.*** p<0.01, ** p<0.05, * p<0.1. The regression with college graduation as outcome only considers respondents who were 26 and over at the time of their last interview. Similarly the cutoff is 22 for college attendance. All races are that of the child. In every interview the CNLSY asks its respondents their highest educational qualification. There are correspondents who mention that they are still in college and their highest qualification till date is college attendance. A respondent is classified as premature if he/she was born early or late.

APPENDIX C

CHAPTER 3 APPENDIX

Table C.1 Summary Stats for Home Owners and Renters

	(1)	(2)	(3)	(4)
Home Owner's				
VARIABLES	Mean	S.D.	Minimum	Maximum
Mother's Age at Birth	27.01	5.400	13	45
Income ($ 10,000)	9.602	13.097	0	146.115
Mother's Education at Birth	12.94	2.445	0	20
Private School	0.138	0.344	0	1
Special Education	0.0915	0.288	0	1
Catholic	0.393	0.489	0	1
Non-Native English Speaker	0.0346	0.183	0	1
Hispanic	0.191	0.393	0	1
Black	0.182	0.386	0	1
4-Year Home Price Change($100,000)	0.444	1.316	-23.23	23.83
Renter's				
Mother's Age at Birth	23.99	5.518	13	45
Income ($ 10,000)	3.695	5.756	0	146.115
Mother's Education at Birth	11.42	2.055	0	20
Private School	0.0508	0.220	0	1
Special Ed.	0.149	0.356	0	1
Catholic	0.430	0.495	0	1
Native English Speaker	0.0677	0.251	0	1
Hispanic	0.237	0.425	0	1
Black	0.489	0.500	0	1

Note: For home owners, the number of observations are 10,505 , and for renters it is 7,712. In panel A, all home values are self-reported and apply to the home owner. All monetary means are in real 2010 dollars and were inflated using the CPI-U.

Table C.2 Housing price and change in school

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Private-Public		Private-Private		Public-Private	
Home value change	0.0066***	0.0066***	0.0021	0.0021	0.0012	0.0012
	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)
Missing Income Dummy	0.0030	0.0030	0.0325***	0.0325**	0.0104***	0.0104**
	(0.007)	(0.008)	(0.008)	(0.013)	(0.004)	(0.005)
Catholic	0.0043	0.0043	0.0277***	0.0277***	0.0076**	0.0076
	(0.005)	(0.009)	(0.006)	(0.010)	(0.003)	(0.005)
Male	-0.0056	-0.0056	-0.0073	-0.0073	0.0061*	0.0061
	(0.005)	(0.006)	(0.006)	(0.009)	(0.003)	(0.004)
Mother's Age at Birth	0.0047***	0.0047***	0.0070***	0.0070***	0.0011***	0.0011***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
BirthOrder	-0.0120***	-0.0120**	-0.0206***	-0.0206***	-0.0070***	-0.0070**
	(0.003)	(0.005)	(0.004)	(0.006)	(0.002)	(0.003)
Family income	0.0039**	0.0039**	0.0124***	0.0124***	0.0027***	0.0027***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.001)	(0.001)
Black	-0.0010	-0.0010	-0.0544***	-0.0544***	0.0067*	0.0067
	(0.006)	(0.009)	(0.009)	(0.012)	(0.004)	(0.007)
Hispanic	-0.0067	-0.0067	-0.0286***	-0.0286***	0.0033	0.0033
	(0.006)	(0.009)	(0.008)	(0.010)	(0.004)	(0.005)
SE Clustered	NO	YES	NO	YES	NO	YES
Observations	7,386	7,386	7,386	7,386	7,386	7,386

Note: Odd numbered columns have robust standard errors and even columns have standard errors clustered at the region level. *** p<0.01, ** p<0.05, * p<0.1. All results in the table come from one multinomial logit model and include home owners only. Housing price changes are real housing price changes between first reported school year and last reported school year. The omitted category is staying in public school. Housing price changes and family income are for every $100,000.

Table C.3 Housing price and change in school, by Catholicity

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Private-Public		Private-Private		Public-Private	
Δ HomeValueXCatholic	0.0039**	0.0039	0.0006	0.0006	0.0022	0.0022
	(0.002)	(0.004)	(0.003)	(0.005)	(0.001)	(0.002)
ΔHomeValueXNon-Catholic	0.0087***	0.0087**	0.0036	0.0036	-0.0010	-0.0010
	(0.002)	(0.004)	(0.003)	(0.005)	(0.003)	(0.005)
Missing Income Dummy	0.0026	0.0026	0.0323***	0.0323**	0.0106***	0.0106**
	(0.007)	(0.008)	(0.008)	(0.013)	(0.004)	(0.005)
Catholic	0.0074	0.0074	0.0286***	0.0286***	0.0062*	0.0062
	(0.005)	(0.009)	(0.006)	(0.008)	(0.004)	(0.006)
Male	-0.0057	-0.0057	-0.0073	-0.0073	0.0060*	0.0060
	(0.005)	(0.007)	(0.006)	(0.009)	(0.003)	(0.004)
Mother's Age at Birth	0.0048***	0.0048***	0.0070***	0.0070***	0.0011***	0.0011***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
BirthOrder	-0.0122***	-0.0122**	-0.0207***	-0.0207***	-0.0069***	-0.0069**
	(0.003)	(0.005)	(0.004)	(0.006)	(0.002)	(0.003)
Family income	0.0039**	0.0039**	0.0124***	0.0124***	0.0027***	0.0027***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Black	-0.0008	-0.0008	-0.0544***	-0.0544***	0.0066*	0.0066
	(0.006)	(0.009)	(0.009)	(0.012)	(0.004)	(0.007)
Hispanic	-0.0069	-0.0069	-0.0285***	-0.0285***	0.0033	0.0033
	(0.006)	(0.009)	(0.008)	(0.009)	(0.004)	(0.005)
SE Clustered	NO	YES	NO	YES	NO	YES
Observations	7,386	7,386	7,386	7,386	7,386	7,386

Note: Odd numbered columns have robust standard errors and even columns have standard errors clustered at the region level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. All results in the table come from one multinomial logit model and include home owners only. Housing price changes are real housing price changes between first reported school year and last reported school year. The omitted category is staying in public school. Housing price changes and family income are for every \$100,000.

Table C.4 Housing price and change in school, by income

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Private-Public		Private-Private		Public-Private	
ΔHomeValueX1stQuartile	0.0105***	0.0105	0.0099	0.0099	-0.0002	-0.0002
	(0.004)	(0.007)	(0.006)	(0.011)	(0.002)	(0.004)
ΔHomeValueX2ndQuartile	0.0040	0.0040	-0.0076	-0.0076	-0.0141***	-0.0141**
	(0.006)	(0.010)	(0.014)	(0.026)	(0.004)	(0.006)
ΔHomeValueX3rdQuartile	-0.0002	-0.0002	0.0072	0.0072	0.0028	0.0028
	(0.006)	(0.011)	(0.007)	(0.013)	(0.006)	(0.008)
ΔHomeValueX4thQuartile	0.0059***	0.0059***	-0.0005	-0.0005	0.0012	0.0012
	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
SecondQuartile	-0.0031	-0.0031	-0.0642***	-0.0642***	-0.0128**	-0.0128
	(0.009)	(0.015)	(0.013)	(0.020)	(0.005)	(0.008)
ThirdQuartile	-0.0021	-0.0021	-0.0138	-0.0138	-0.0119**	-0.0119
	(0.008)	(0.011)	(0.010)	(0.014)	(0.005)	(0.007)
FourthQuartile	0.0190***	0.0190*	0.0281***	0.0281*	0.0009	0.0009
	(0.007)	(0.011)	(0.009)	(0.015)	(0.004)	(0.006)
SE Clustered	NO	YES	NO	YES	NO	YES
Observations	7,386	7,386	7,386	7,386	7,386	7,386

Note: Odd numbered columns have robust standard errors and even columns have standard errors clustered at the region level. *** p<0.01, ** p<0.05, * p<0.1. All results in the table come from one multinomial logit model and include home owners only. Housing price changes are real housing price changes between first reported school year and last reported school year. The omitted category is staying in public school. All specifications control for Catholicity, gender, mothers age at birth, birth order, and race. Housing price changes and family income are for every $100,000.

Table C.5 Housing price and change in school, by gender

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Private-Public		Private-Private		Public-Private	
ΔHomeValueXBoy	0.0077***	0.0077**	0.0049	0.0049	-0.0016	-0.0016
	(0.002)	(0.003)	(0.003)	(0.005)	(0.001)	(0.001)
ΔHomeValueXGirl	0.0059***	0.0059	-0.0006	-0.0006	0.0031**	0.0031
	(0.002)	(0.004)	(0.003)	(0.005)	(0.002)	(0.004)
Male	-0.0067	-0.0067	-0.0097	-0.0097	0.0079**	0.0079**
	(0.005)	(0.006)	(0.006)	(0.009)	(0.003)	(0.004)
Missing Income Dummy	0.0030	0.0030	0.0324***	0.0324**	0.0104***	0.0104**
	(0.007)	(0.008)	(0.008)	(0.013)	(0.004)	(0.005)
Catholic	0.0043	0.0043	0.0276***	0.0276***	0.0075**	0.0075
	(0.005)	(0.009)	(0.006)	(0.008)	(0.003)	(0.005)
Mother's Age at Birth	0.0047***	0.0047***	0.0070***	0.0070***	0.0011***	0.0011***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
BirthOrder	-0.0120***	-0.0120**	-0.0207***	-0.0207***	-0.0069***	-0.0069**
	(0.003)	(0.005)	(0.004)	(0.006)	(0.002)	(0.003)
Family income	0.0039**	0.0039**	0.0124***	0.0124***	0.0027***	0.0027***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Black	-0.0009	-0.0009	-0.0543***	-0.0543***	0.0065*	0.0065
	(0.006)	(0.009)	(0.009)	(0.012)	(0.004)	(0.007)
Hispanic	-0.0065	-0.0065	-0.0284***	-0.0284***	0.0031	0.0031
	(0.006)	(0.009)	(0.008)	(0.010)	(0.004)	(0.005)
SE Clustered	NO	YES	NO	YES	NO	YES
Observations	7,386	7,386	7,386	7,386	7,386	7,386

Note: Odd numbered columns have robust standard errors and even columns have standard errors clustered at the region level. *** p<0.01, ** p<0.05, * p<0.1. All results in the table come from one multinomial logit model and include home owners only. Housing price changes are real housing price changes between first reported school year and last reported school year. The omitted category is staying in public school. Housing price changes and family income are for every $100,000.

Table C.6 Housing price and change in school, by race

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Private-Public		Private-Private		Public-Private	
ΔHomeValueXBlack	0.0073	0.0073	0.0134	0.0134	-0.0068*	-0.0068
	(0.008)	(0.015)	(0.012)	(0.022)	(0.004)	(0.006)
ΔHomeValueXHispanic	0.0081***	0.0081**	0.0007	0.0007	-0.0021	-0.0021
	(0.002)	(0.004)	(0.004)	(0.007)	(0.003)	(0.003)
ΔHomeValueXWhite	0.0061***	0.0061	0.0020	0.0020	0.0024	0.0024
	(0.002)	(0.004)	(0.003)	(0.005)	(0.001)	(0.002)
Black	-0.0015	-0.0015	-0.0565***	-0.0565***	0.0080**	0.0080
	(0.006)	(0.009)	(0.010)	(0.013)	(0.004)	(0.007)
Hispanic	-0.0084	-0.0084	-0.0281***	-0.0281***	0.0050	0.0050
	(0.007)	(0.011)	(0.008)	(0.009)	(0.004)	(0.005)
Missing Income Dummy	0.0030	0.0030	0.0322***	0.0322**	0.0104***	0.0104**
	(0.007)	(0.008)	(0.008)	(0.013)	(0.004)	(0.005)
Catholic	0.0044	0.0044	0.0279***	0.0279***	0.0073**	0.0073
	(0.005)	(0.009)	(0.006)	(0.008)	(0.003)	(0.005)
Male	-0.0055	-0.0055	-0.0072	-0.0072	0.0058*	0.0058
	(0.005)	(0.006)	(0.006)	(0.009)	(0.003)	(0.004)
Mother's Age at Birth	0.0047***	0.0047***	0.0070***	0.0070***	0.0011***	0.0011***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
BirthOrder	-0.0120***	-0.0120**	-0.0207***	-0.0207***	-0.0069***	-0.0069**
	(0.003)	(0.005)	(0.004)	(0.006)	(0.002)	(0.003)
Family income	0.0040**	0.0040***	0.0124***	0.0124***	0.0027***	0.0027***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
SE Clustered	NO	YES	NO	YES	NO	YES
Observations	7,386	7,386	7,386	7,386	7,386	7,386

Note: Odd numbered columns have robust standard errors and even columns have standard errors clustered at the region level. *** p<0.01, ** p<0.05, * p<0.1. All results in the table come from one multinomial logit model and include home owners only. Housing price changes are real housing price changes between first reported school year and last reported school year. The omitted category is staying in public school. Housing price changes and family income are for every $100,000.

Table C.7 Estimates for transition from middle to high school

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Private-Public		Private-Private		Public-Private	
Home value change	0.0086	0.0086	0.0015	0.0015	0.0090**	0.0090
	(0.006)	(0.008)	(0.008)	(0.013)	(0.004)	(0.008)
Black	0.0026	0.0026	-0.0196*	-0.0196	0.0035	0.0035
	(0.008)	(0.013)	(0.011)	(0.013)	(0.007)	(0.009)
Hispanic	0.0027	0.0027	-0.0149	-0.0149	0.0100	0.0100
	(0.010)	(0.013)	(0.012)	(0.017)	(0.007)	(0.007)
Family income	0.0000	0.0000	0.0000***	0.0000***	0.0000	0.0000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing Income Dummy	0.0127	0.0127	-0.0027	-0.0027	0.0086	0.0086
	(0.008)	(0.008)	(0.013)	(0.016)	(0.006)	(0.011)
Catholic	0.0007	0.0007	0.0300***	0.0300**	-0.0025	-0.0025
	(0.007)	(0.008)	(0.010)	(0.014)	(0.006)	(0.008)
Male	-0.0042	-0.0042	0.0086	0.0086	0.0058	0.0058
	(0.007)	(0.011)	(0.009)	(0.009)	(0.006)	(0.009)
Mother's Age at Birth	0.0018	0.0018	0.0072***	0.0072***	-0.0008	-0.0008
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
BirthOrder	-0.0051	-0.0051	-0.0366***	-0.0366***	-0.0003	-0.0003
	(0.005)	(0.007)	(0.008)	(0.009)	(0.004)	(0.005)
SE Clustered	NO	YES	NO	YES	NO	YES
Observations	1,959	1,959	1,959	1,959	1,959	1,959

Note: Odd numbered columns have robust standard errors and even columns have standard errors clustered at the region level. *** p<0.01, ** p<0.05, * p<0.1. All results in the table come from one multinomial logit model and include home owners only. Housing price changes are real housing price changes between first reported middle school year and first reported high school year. The omitted category is staying in public school. Housing price changes and family income are for every $100,000.

Table C.8 Estimates for transition from primary to middle school

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Private-Public		Private-Private		Public-Private	
Home value change	0.0056***	0.0056	0.0018	0.0018	0.0004	0.0004
	(0.002)	(0.004)	(0.003)	(0.004)	(0.002)	(0.004)
Black	-0.0025	-0.0025	-0.0506***	-0.0506***	0.0026	0.0026
	(0.009)	(0.014)	(0.015)	(0.021)	(0.006)	(0.009)
Hispanic	0.0044	0.0044	-0.0426***	-0.0426*	0.0032	0.0032
	(0.009)	(0.016)	(0.014)	(0.022)	(0.006)	(0.008)
Family income	0.0000**	0.0000	0.0000***	0.0000*	0.0000***	0.0000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing Income Dummy	0.0174*	0.0174	0.0172	0.0172	0.0013	0.0013
	(0.010)	(0.013)	(0.015)	(0.020)	(0.007)	(0.009)
Catholic	0.0149**	0.0149	0.0446***	0.0446***	0.0068	0.0068
	(0.007)	(0.011)	(0.010)	(0.010)	(0.005)	(0.006)
Male	0.0005	0.0005	-0.0104	-0.0104	0.0017	0.0017
	(0.007)	(0.012)	(0.009)	(0.015)	(0.005)	(0.007)
Mother's Age at Birth	0.0022***	0.0022***	0.0064***	0.0064***	0.0010*	0.0010***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.000)
BirthOrder	-0.0091***	-0.0091	-0.0182***	-0.0182***	-0.0083**	-0.0083**
	(0.003)	(0.006)	(0.005)	(0.007)	(0.003)	(0.004)
SE Clustered	NO	YES	NO	YES	NO	YES
Observations	3,240	3,240	3,240	3,240	3,240	3,240

Note: Odd numbered columns have robust standard errors and even columns have standard errors clustered at the region level. *** p<0.01, ** p<0.05, * p<0.1. All results in the table come from one multinomial logit model and include home owners only. Housing price changes are real housing price changes between first reported primary school year and first reported middle school year. The omitted category is staying in public school. Housing price changes and family income are for every $100,000.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abadie, A., A. Diamond, and J. Hainmueller (2010, June). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association 105*(490), 493–505.

Abadie, A. and J. Gardeazabal (2003, February). The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review 93*(1), 113–132.

Abraham, S. and L. Sun (2018). Estimating Dynamic Treatment Effects in Event Studies With Heterogeneous Treatment Effects. *SSRN Electronic Journal*.

Ahn, T. and A. Yelowitz (2015, October). The short-run impacts of Connecticut's paid sick leave legislation. *Applied Economics Letters 22*(15), 1267–1272.

Ahn, T. and A. Yelowitz (2016). Paid Sick Leave and Absenteeism: The First Evidence from the U.S. *SSRN Electronic Journal*.

Allison, P. D. and F. F. Furstenberg (1989). How marital dissolution affects children: Variations by age and sex. *Developmental Psychology 25*(4), 540–549.

Altonji, J., T. Elder, and C. Taber (2005, February). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy 113*(1), 151–184.

Antwi, Y. A. and J. R. Bowblis (2018, May). The Impact of Nurse Turnover on Quality of Care and Mortality in Nursing Homes: Evidence from the Great Recession. *American Journal of Health Economics 4*(2), 131–163.

Appelbaum, E. and H. Jorgensen (2014). Expanding Federal Family and Medical Leave Coverage: Who Benefits from Changes in Eligibility Requirements?

Applebaum, E. and R. Milkman (2016). No big deal: The impact of New York City's paid sick days law on employers.

Associates Inc, A. (2013). Nursing Home Compare Five-Star Quality Rating System.

Athey, S. and G. W. Imbens (2017, May). The State of Applied Econometrics: Causality and Policy Evaluation. *Journal of Economic Perspectives 31*(2), 3–32.

Autor, D., D. Figlio, K. Karbownik, J. Roth, and M. Wasserman (2019, July). Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes. *American Economic Journal: Applied Economics 11*(3), 338–381.

Autor, D. and M. Wasserman (2013). Wayward Sons.

Baker, M. and K. Milligan (2016). Boy-Girl Differences in Parental Time Investments: Evidence from Three Countries. *Journal of Human Capital 10*(4).

Barnard-Brak, L., X. Liu, and T. Wei (2015). Gender Differences in Mathematics and Reading Trajectories among Children from Kindergarten to Eighth Grade. *Research in Education 93*(1).

Barrow, L. (2006, December). Private school location and neighborhood characteristics. *Economics of Education Review 25*(6), 633–645.

Benjamin, J. D., P. Chinloy, and G. D. Jud (2004, November). Real Estate Versus Financial Wealth in Consumption. *The Journal of Real Estate Finance and Economics 29*(3), 341–354.

Bennett, P., R. Peach, and S. Peristiani (2001, November). Structural Change in the Mortgage Market and the Propensity to Refinance. *Journal of Money, Credit and Banking 33*(4), 955.

Bertrand, M. and J. Pan (2013, January). The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. *American Economic Journal: Applied Economics 5*(1), 32–64.

Betts, J. R. and R. W. Fairlie (2003, May). Does immigration induce ‘native flight’ from public schools into private schools? *Journal of Public Economics 87*(5-6), 987–1012.

Borusyak, K. and X. Jaravel (2016). Revisiting Event Study Designs. *SSRN Electronic Journal*.

Brandeis, G. H., W. L. Ooi, M. Hossain, J. N. Morris, and L. A. Lipsitz (1994, April). A Longitudinal Study of Risk Factors Associated with the Formation of Pressure Ulcers in Nursing Homes. *Journal of the American Geriatrics Society 42*(4), 388–393.

Broughman, S. and N. Swain (2016). Characteristics of Private Schools in the United States: Results From the 2013-14 Private School Universe Survey.

Buchmann, C. and T. A. DiPrete (2006). The growing female advantage in college completion: The role of family background and academic achievement. *American Sociological Review 71*(4).

Buddin, R. J., J. J. Cordes, and S. N. Kirby (1998, July). School Choice in California: Who Chooses Private Schools? *Journal of Urban Economics 44*(1), 110–134.

Campbell, J. Y. and J. F. Cocco (2007, April). How do house prices affect consumption?

Evidence from micro data. *Journal of Monetary Economics 54*(3), 591–621.

Castle, N. G. and J. C. Ferguson (2010, August). What Is Nursing Home Quality and How Is It Measured? *The Gerontologist 50*(4), 426–442.

Cawley, J., D. C. Grabowski, and R. A. Hirth (2006, March). Factor substitution in nursing homes. *Journal of Health Economics 25*(2), 234–247.

CDC (2009). Guideline for Prevention of Catheter-Associated Urinary Tract Infections.

Chakrabarti, R. and J. Roy (2016, January). Do charter schools crowd out private school enrollment? Evidence from Michigan. *Journal of Urban Economics 91*, 88–103.

Charles, K. K., E. Hurst, and M. J. Notowidigdo (2018, October). Housing Booms and Busts, Labor Market Opportunities, and College Attendance. *American Economic Review 108*(10), 2947–2994.

Chen, J. (2019). *Essays on Occupational Regulation and Paid Sick Leave.* Ph. D. thesis, Lehigh University.

Chen, M. M. and D. C. Grabowski (2015, July). Intended and Unintended Consequences of Minimum Staffing Standards for Nursing Homes: NURSING HOME STAFFING STANDARDS. *Health Economics 24*(7), 822–839.

Cherlin, A. J., F. F. Furstenberg, P. L. Chase-Lansdale, K. E. Kiernan, P. K. Robins, D. R. Morrison, and J. O. Teitler (1991, June). Longitudinal Studies of Effects of Divorce on Children in Great Britain and the United States. *Science 252*(5011), 1386–1389.

Chetty, R., N. Hendren, F. Lin, J. Majerovitz, and B. Scuderi (2016, May). Childhood Environment and Gender Gaps in Adulthood. *American Economic Review 106*(5), 282–288.

Cobb-Clark, D. A. and E. Tekin (2014, June). Fathers and youths' delinquent behavior. *Review of Economics of the Household 12*(2), 327–358.

Cohen-Zada, D. and T. Elder (2009, July). Historical religious concentrations and the effects of Catholic schooling. *Journal of Urban Economics 66*(1), 65–74.

Collins, A. and M. C. Snell (2000, June). Parental preferences and choice of school. *Applied Economics 32*(7), 803–813.

Dahl, G. B. and E. Moretti (2008, October). The Demand for Sons. *Review of Economic Studies 75*(4), 1085–1120.

Davis, J. and H. Renski (2020, October). Do Industrial Preservation Policies Protect and

Promote Urban Industrial Activity?: Examining the Impact of New York City's Industrial Business Zone Program. *Journal of the American Planning Association 86*(4), 431–442.

Davis M., A. (2011). *Why Do Parents Choose to Send Their Children To Private Schools?* Ph. D. thesis, Georgia Southern University.

de Chaisemartin, C. and X. D'Haultfœuille (2020, September). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review 110*(9), 2964–2996.

Deep, A. and D. Domanski (2002). Housing markets and economic growth: lessons from the US refinancing boom.

DeRigne, L., P. Stoddard-Dare, and L. Quinn (2016, March). Workers Without Paid Sick Leave Less Likely To Take Time Off For Illness Or Injury Compared To Those With Paid Sick Leave. *Health Affairs 35*(3), 520–527.

Dill, J. S., J. C. Morgan, and V. W. Marshall (2013, April). Contingency, Employment Intentions, and Retention of Vulnerable Low-wage Workers: An Examination of Nursing Assistants in Nursing Homes. *The Gerontologist 53*(2), 222–234.

Dorr, D. A., S. D. Horn, and R. J. Smout (2005, May). Cost Analysis of Nursing Home Registered Nurse Staffing Times: COST ANALYSIS OF NH NURSE STAFFING TIMES. *Journal of the American Geriatrics Society 53*(5), 840–845.

Drago, R. and V. Lowell (2011). San Francisco's Paid Sick Leave Ordinance: Outcomes for Employers and Employees.

Dube, A., T. W. Lester, and M. Reich (2016, July). Minimum Wage Shocks, Employment Flows, and Labor Market Frictions. *Journal of Labor Economics 34*(3), 663–704.

Dube, A. and B. Zipperer (2015). Pooling Multiple Case Studies Using Synthetic Controls: An Application to Minimum Wage Policies. *SSRN Electronic Journal*.

Elder, T. and C. Jepsen (2014, March). Are Catholic primary schools more effective than public primary schools? *Journal of Urban Economics 80*, 28–38.

Fairlie, R. W. and A. M. Resch (2002, February). Is There "White Flight" into Private Schools? Evidence from the National Educational Longitudinal Survey. *Review of Economics and Statistics 84*(1), 21–33.

Fan, X., H. Fang, and S. Markussen (2015). Mothers' Employment and Children's Educational Gender Gap.

Favreault, M. and J. Dey (2016). Long-Term Services and Supports for Older Americans:

Risks and Financing Research Brief.

Federal Reserve Bank of San Francisco, M. Doms, J. Krainer, and Federal Reserve Bank of San Francisco (2007, July). Innovations in Mortgage Markets and Increased Spending on Housing. *Federal Reserve Bank of San Francisco, Working Paper Series*, 1.000–43.000.

for Health, N. I. (2019).

for Medicaid and Medicare Services, C. (2015). Nursing Home Data Compendium.

for Medicaid and Medicare Services, C. (2019). Prospective Payment Systems - General Information.

for Medicare and Medicaid, C. (2020). Nursing Home Compare Data Archive.

Gertler, P. J. (1989, February). Subsidies, quality, and the regulation of nursing homes. *Journal of Public Economics 38*(1), 33–52.

Goetzel, R. Z., S. R. Long, R. J. Ozminkowski, K. Hawkins, S. Wang, and W. Lynch (2004, April). Health, Absence, Disability, and Presenteeism Cost Estimates of Certain Physical and Mental Health Conditions Affecting U.S. Employers:. *Journal of Occupational and Environmental Medicine 46*(4), 398–412.

Goldring, E. B. and K. J. Phillips (2008, May). Parent preferences and parent choices: the public–private decision about school choice. *Journal of Education Policy 23*(3), 209–230.

Goodman-Bacon, A. (2021, December). Difference-in-differences with variation in treatment timing. *Journal of Econometrics 225*(2), 254–277.

Grabowski, D. C., J. R. Bowblis, J. A. Lucas, and S. Crystal (2011, July). Labor Prices and the Treatment of Nursing Home Residents with Dementia. *International Journal of the Economics of Business 18*(2), 273–292.

Greenspan, A. and J. E. Kennedy (2005). Estimates of Home Mortgage Originations, Repayments, and Debt on One-to-Four-Family Residences. *SSRN Electronic Journal*.

Harrington, C., S. Woolhandler, J. Mullan, H. Carrillo, and D. U. Himmelstein (2002, April). Does Investor-Ownership of Nursing Homes Compromise the Quality of Care? *International Journal of Health Services 32*(2), 315–325.

Harrington, C., D. Zimmerman, S. L. Karon, J. Robinson, and P. Beutel (2000, September). Nursing Home Staffing and Its Relationship to Deficiencies. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences 55*(5), S278–S287.

Houser, A., W. Fox-Grage, and K. Ujvari (2018). Across the States 2018: Profiles of Long-

Term Services and Supports.

Huang, S. S. and J. R. Bowblis (2019, November). Is the Quality of Nursing Homes Countercyclical? Evidence From 2001 Through 2015. *Gerontologist 59(6)*, 1044–1054.

Jacobson, L. S., R. J. Lalonde, and D. G. Sullivan (1993, September). Earnings Losses of Displaced Workers. *The American Economic Review 83*(4), 685–709.

Jepsen, C. (2003). The Effectiveness of Catholic Primary Schooling. *Journal of Human Resources XXXVIII*(4), 928–941.

Kalil, A., M. Mogstad, M. Rege, and M. E. Votruba (2016, October). Father Presence and the Intergenerational Transmission of Educational Attainment. *Journal of Human Resources 51*(4), 869–899.

Keith, V. M. and B. Finlay (1988, August). The Impact of Parental Divorce on Children's Educational Attainment, Marital Timing, and Likelihood of Divorce. *Journal of Marriage and the Family 50*(3), 797.

Krein, S. F. and A. H. Beller (1988, May). Educational attainment of children from single-parent families: Differences by exposure, gender, and race. *Demography 25*(2), 221–234.

Kristoffersen, J. H., C. Obel, and N. Smith (2015, July). Gender differences in behavioral problems and school outcomes. *Journal of Economic Behavior & Organization 115*, 75–93.

Li, M. (2009, June). Is there "white flight" into private schools? New evidence from High School and Beyond. *Economics of Education Review 28*(3), 382–392.

Lindemann, D. and D. Britton (2015). Earned Sick Days in Jersey City : A Study of Employers and Employees at Year One.

Long, J. E. and E. F. Toma (1988, May). The Determinants of Private School Attendance, 1970-1980. *The Review of Economics and Statistics 70*(2), 351.

Lovenheim, M. and K. J. Mumford (2010). Do Family Wealth Shocks Affect Fertility Choices? Evidence from the Housing Market Boom and Bust. *SSRN Electronic Journal*.

Lovenheim, M. F. (2011, October). The Effect of Liquid Housing Wealth on College Enrollment. *Journal of Labor Economics 29*(4), 741–771.

Lovenheim, M. F. and C. L. Reynolds (2013). The Effect of Housing Wealth on College Choice: Evidence from the Housing Boom. *Journal of Human Resources 48*(1), 1–35.

Matsudaira, J. D. (2014, March). Monopsony in the Low-Wage Labor Market? Evidence from Minimum Nurse Staffing Regulations. *The Review of Economics and Statistics 96*(1), 92–102.

Murnane, R. J. and S. F. Reardon (2018, February). Long-Term Trends in Private School Enrollments by Family Income. *AERA Open 4*(1), 233285841775135.

OECD (2013). Education at a Glance.

of Kentucky, U. (2020). National Welfare Data.

of Labor Statistics, B. (2020a).

of Labor Statistics, B. (2020b). Local Area Unemployment Statistics.

of Medicine, I. (1986). Improving the quality of care in nursing homes.

of Public Health, B. S. (2020). Long-Term Care Focus.

of State Legislatures, N. C. (2019).

ONET (2018). publisher: Onet Online.

Owens, J. (2016, July). Early Childhood Behavior Problems and the Gender Gap in Educational Attainment in the United States. *Sociology of Education 89*(3), 236–258.

Pandey, L., D. L. Sjoquist, and M. B. Walker (2009, January). An Analysis of Private School Closings. *Education Finance and Policy 4*(1), 34–59.

Park, J. and S. C. Stearns (2009, February). Effects of State Minimum Staffing Standards on Nursing Home Staffing and Quality of Care. *Health Services Research 44*(1), 56–78.

Pichler, S. and N. R. Ziebarth (2019, March). Reprint of: The pros and cons of sick pay schemes: Testing for contagious presenteeism and noncontagious absenteeism behavior. *Journal of Public Economics 171*, 86–104.

Prevoo, T. and B. ter Weel (2014). The Effect of Family Disruption on Children's Personality Development: Evidence from British Longitudinal Data. *SSRN Electronic Journal*.

Reardon, S. F., J. E. Cheadle, and J. P. Robinson (2009, January). The Effect of Catholic Schooling on Math and Reading Development in Kindergarten Through Fifth Grade. *Journal of Research on Educational Effectiveness 2*(1), 45–87.

Reports, N. V. S. (2018). Deaths: Final data for 2018.

Rodgers, J. L., W. H. Beasley, D. E. Bard, K. M. Meredith, M. D. Hunter, A. B. Johnson, M. Buster, C. Li, K. O. May, S. Mason Garrison, W. B. Miller, E. van den Oord, and

D. C. Rowe (2016, July). The NLSY Kinship Links: Using the NLSY79 and NLSY-Children Data to Conduct Genetically-Informed and Family-Oriented Research. *Behavior Genetics 46*(4), 538–551.

Ruffini, K. (2021). Worker Earnings, Service Quality, and Firm Profitability: Evidence from Nursing Homes and Minimum Wage Reforms. *SSRN Electronic Journal*.

S. Jacobson, L., R. J. LaLonde, and D. G. Sullivan (1993, September). Earnings Losses of Displaced Workers. *The American Economic Review 83*(4), 685–709.

Saint, S. (2000a). Clinical and economic consequences of nosocomial catheter-related bacteriuria.

Saint, S. (2000b, February). Clinical and economic consequences of nosocomial catheter-related bacteriuria. *American Journal of Infection Control 28*(1), 68–75.

Sojourner, A. J., B. R. Frandsen, R. J. Town, D. C. Grabowski, and M. M. Chen (2015, August). Impacts of Unionization on Quality and Productivity: Regression Discontinuity Evidence from Nursing Homes. *ILR Review 68*(4), 771–806.

Squillace, M. R., R. E. Remsburg, L. D. Harris-Kojetin, A. Bercovitz, E. Rosenoff, and B. Han (2009, April). The National Nursing Assistant Survey: Improving the Evidence Base for Policy Initiatives to Strengthen the Certified Nursing Assistant Workforce. *The Gerontologist 49*(2), 185–197.

Stearns, J. and C. White (2016). Can Paid Sick Leave Mandates Reduce Leave-Taking? *SSRN Electronic Journal*.

Stevens, A. H., D. L. Miller, M. E. Page, and M. Filipski (2015, November). The Best of Times, the Worst of Times: Understanding Pro-cyclical Mortality. *American Economic Journal: Economic Policy 7*(4), 279–311.

Tominey, E. (2016, August). Female labour supply and household employment shocks: Maternity leave as an insurance mechanism. *European Economic Review 87*, 256–271.

Zhao, L. and G. Burge (2017, January). Housing Wealth, Property Taxes, and Labor Supply among the Elderly. *Journal of Labor Economics 35*(1), 227–263.