SciencesPo
SCHOOL OF PUBLIC AFFAIRS

PUBLIC POLICY MASTER THESIS

April 2026

The Political Economy of Kindergarten Expansion

Eve Margolis

Master's Thesis supervised by Silvia Vannutelli

Second member of the jury: Roberto Galbiati

Master in Public Policy

Policy Stream: Economics and Public Policy

Acknowledgements

I would like to express my gratitude to all of the people who supported and guided me in the completion of this Master's thesis. First, I warmly thank my thesis supervisor, Professor Silvia Vannutelli, for her guidance and insightful advice. I am also deeply grateful to Professor Roberto Galbiati for having agreed to be the second member of my jury.

I would also like to thank my family, loved ones, and friends for their unwavering support during these intense months.

Contents

Why should I read this research?

Whether expanding public goods generates electoral returns for incumbents is a central question in political economy. This thesis uses the staggered adoption of full-day kindergarten (FDK) across Pennsylvania school districts to study the political consequences of a locally delivered education policy, exploiting quasi-experimental variation in the timing of adoption via the Callaway and Sant'Anna (2021) doubly-robust difference-in-differences estimator. The analysis merges individual-level records from the Pennsylvania SURE voter file (8.9 million registrants) with Department of Education enrollment data (2008–2025), American Community Survey estimates, and state legislative returns.

Three main findings emerge. First, FDK adoption increases new voter registration rates among adults of plausible parenting age by approximately 0.33 percentage points per year. This represents a relative increase of 20 to 35 percent above the pre-treatment baseline. Second, this effect is present for both mothers and fathers in similar magnitudes (0.30 and 0.36 percentage point increases respectively), builds gradually, and reaches statistical significance only after nine to ten years under simultaneous confidence bands. It generates no detectable net partisan advantage in new registrations. This pattern points toward a diffuse civic-mobilization channel rather than direct credit-claiming by incumbents. Third, single mothers of kindergarten-age children display a positive but statistically imprecise increase in labor force participation, suggestive of a time-constraint mechanism, while two-parent households show a null or negative response. Electoral findings are more ambiguous. The thesis combines two distinct empirical exercises. A descriptive national state-level analysis suggests positive associations between kindergarten expansion and some Democratic outcomes. A separate causal district-level design in Pennsylvania instead identifies a robust registration effect but mixed and often negative Democratic electoral coefficients, likely reflecting non-partisan mobilization and boundary-mapping limitations.

Several limitations qualify the scope of these claims. The electoral outcome estimates suffer primarily from the mismatch between school-district and legislative-district boundaries, which attenuates and confounds mapping between treatment and vote shares. The causal design is restricted to Pennsylvania, and treatment timing is recovered from administrative enrollment data. The estimates apply specifically to the 142 treated districts that transitioned to FDK during the sample window. Finally, the registration effect only materializes after approximately a decade, beyond the horizon of a single electoral term.

The contribution rests on three concrete design features: the use of individual-level voter-file records rather than self-reported survey measures, a modern staggered difference-in-differences estimator applied to a panel of school districts, and the direct link from district-level policy variation to individual registration outcomes. This positive and statistically significant relationship holds over a complete grid of specifications, suggesting robust results. The findings suggest that FDK expansion is a civically meaningful policy whose political payoff accrues slowly and non-partisanly, a coherent result with direct implications for how educational subsidies should be targeted and communicated to local decision-makers.

1. Introduction

1.1 Puzzle, motivation, and scope

1.1.1 Context and motivation

In the United States there is no federal policy regarding kindergarten attendance. Instead, each state is responsible for its own policies, which results in only 21 states having mandatory, full-day kindergarten for every child above the age of 5 as of 2020. The remaining 29 states implement different types of policies or, in some cases, none. Some states delegate the decision concerning mandatory attendance to smaller geographical entities, such as school districts or local educational agencies.

Whether full-day kindergarten (as opposed to half-day kindergarten or no kindergarten at all) has long-lasting impacts on the children who attend it is a contentious, much-studied topic. What is less often studied is why lawmakers would push for this type of attendance policy. In addition to direct educational or civic benefits, kindergarten has interesting political economy implications. As detailed in Section 3, within a state there are many different elections, including ones for the school boards the local entity most relevant for the governance of US public schools. Free mandatory kindergarten education can be seen as a public good, as it is neither rival nor excludable, and it is moreover very salient, as parents are necessarily made aware of changes in the policies concerning their children's education. Looking at this level of early education from a political economy point of view therefore seems intuitively justified.

1.1.2 Research question

This thesis specifically aims to answer the question "Is it in a politician's electoral interest to expand kindergarten education?" It proceeds in two parts. The first provides a descriptive analysis of the factors that lead states to mandate kindergarten attendance, performing a pseudo intention-to-treat (ITT) analysis at the state level across the entire continental United States. The second studies the causal link between full-day kindergarten provision and political participation among adults who are of an age to plausibly have kindergarten-age children. The policy-relevant outcomes considered include voter registration at the aggregate level and by party, and vote share for incumbents in the election following the switch.

The second part focuses on the state of Pennsylvania, for three reasons. First, Pennsylvania delegates key kindergarten attendance provisions to the local educational agency level, generating rich within-state variation that enables district-level staggered adoption designs. Second, Pennsylvania makes a large volume of administrative data available, including educational enrollment data from the early 2000s to 2025, and a comprehensive individual-level voter registration database, allowing a cleaner link from district-level exposure to subsequent political behavior. Third, Pennsylvania is a politically competitive swing state, making any detected effects more practically relevant.

Two mechanisms are explored. The first concerns FDK as free all-day childcare: by relaxing parental, and particularly maternal, time constraints, FDK may increase community engagement and the propensity to register to vote. The second is a salience channel: because FDK is a visible policy change that is directly experienced by families, expansions may boost support for incumbents associated with the change. Adoption is however likely endogenous, so producing credible identification requires careful analysis of treatment timing. The causal estimates are identified from the 142 districts that transitioned to FDK during the observation window and are not directly generalizable to the 277 always-treated district, whose pre-treatment period is unobserved, nor to the wealthier, larger never-treated districts.

1.1.3 Contributions

This thesis contributes to the political economy literature in three concrete ways. It establishes a causal link between kindergarten expansion and voter registration, showing that a seemingly apolitical policy generates increases of 20 to 35 percent in yearly registration rates among adults of plausible parenting age. It then examines electoral outcomes at the state level in terms of vote margins and seat shares, conditioning on incumbency. The aggregate electoral effects are not straightforwardly positive for the party expanding FDK: Democratic vote shares and margins are if anything negative at both the House and Senate level in the post-adoption period, consistent with non-partisan mobilization and the geographical imprecision of the school-district-to-legislative-district mapping. The one partial exception, a positive and significant Democratic incumbent retention effect at the Senate level, is suggestive of a credit-claiming channel but rests on a small and imperfectly matched sample. Finally, the thesis finds that single mothers of young children experience gradual gains in labor force participation of 3 to 5 percentage points. This provides evidence on a channel through which FDK affects political participation, while suggesting due to the increase in male registration that the effect is broader than a purely maternal time-constraint story would predict.

1.2 Road map

Section 2 provides an overview of the interdisciplinary state of knowledge and demonstrates how this thesis bridges the economics of education and political economics literature. Section 3 explains the institutional context, detailing the kindergarten policy landscape in the US and Pennsylvania and defining how policy changes are studied here. Section 4 describes the data, key variables, and descriptive statistics. Section 5 details the empirical strategies for both parts. Sections 6 and 7 present the main findings, and an additional mechanism check is presented in Section 8. Section 9 addresses threats to identification and performs robustness checks. Section 10 discusses the policy implication, and Section 11 concludes and details the limitations, directions for future research, and policy implications that can be derived from this thesis.

2. Interdisciplinary state of knowledge

This thesis sits at the intersection of two strands of the economics literature that are usually discussed separately. The first studies early education expansions, interpreting changes in schooling time primarily as changes in childcare supply and their downstream effects on parental time allocation. The second is the political economy literature, which studies when politicians choose to expand visible public goods and how voters respond to salient, credit-claimable policies. Since full-day kindergarten functions simultaneously as a family-facing service and a locally delivered public good, it provides a natural setting for bridging these two strands.

Before reviewing each strand in turn, it is worth being precise about where this thesis stands relative to its closest antecedents. Cools (2022) establishes a time-cost channel for political participation by showing that having an infant reduces voter turnout, and Teney et al. (2024) document that parenthood shifts voting behavior in gendered, issue-linked ways. Dhuey et al. (2021) provide the most direct precursor on the supply side, showing that Ontario's full-day kindergarten rollout generated positive intensive-margin labor supply responses among mothers. Gibbs et al. (2025) frames public schooling expansions explicitly as time-allocation shocks, reinforcing the mechanism this thesis investigates. What the present thesis adds to this body of work is not a prior claim of priority but three concrete design features: the use of individual-level voter-file records rather than self-reported survey turnout, a modern staggered difference-in-differences estimator applied to a panel of school districts, and the direct construction of a link from district-level policy variation to individual registration outcomes. These features, taken together, allow for a more credible causal claim than has been possible with the data and designs previously available.

2.1 Early education policies and children's outcomes

The effects of early education on children's outcomes have been studied extensively in the economics of education literature. Although this thesis does not directly focus on these effects, the literature remains relevant for two reasons. It provides context that helps explain why kindergarten expansions are politically controversial and why voters hold strong priors on the topic. It also emphasizes that the benefits of early education policies extend beyond academic performance, which matters when considering downstream effects on civic behavior.

Brilli et al. (2016) find that early childhood policies affect outcomes through multiple channels simultaneously, shaping both mothers' labor market outcomes and children's development. Although the setting, Italy, differs from the United States and involves a different type of kindergarten policy, the central result generalizes: early education policies can generate immediate household-level effects through time and budget constraints, and longer-run effects through children's skills and trajectories. This thesis focuses primarily on the former channel, namely the short-run implications of full-day kindergarten for adults' political behavior, while the child outcome literature provides the policy context that makes these expansions salient and politically meaningful.

2.2 Kindergarten as childcare: parental labor supply and time allocation

A central mechanism in the labor and public economics literature is that schooling time directly affects parental, and most often maternal, time allocation. Moving from no kindergarten or half-day provision to full-day kindergarten relaxes household time constraints, particularly for parents who were not previously enrolled and who therefore had to either purchase childcare or reduce their market work. This channel matters here because a policy that directly alters parents' daily constraints is likely to be both salient and credibly attributed to incumbents.

The empirical evidence is largely consistent with this mechanism. Using Ontario's rollout of full-day kindergarten, Dhuey et al. (2021) find little evidence of extensive-margin labor supply effects but positive intensive-margin responses: mothers of four-year-olds work longer weekly hours and have lower absenteeism rates. This suggests that kindergarten expansions can meaningfully change behavior even without large participation effects. This is precisely the type of margin that could translate into more time for community engagement and administrative tasks such as voter registration. Gibbs et al. (2025) frames public schooling expansions explicitly as changes in the household time endowment, modeling how modifications in schooling time can increase maternal employment and reinforcing the interpretation of full-day kindergarten as an input into household time allocation.

More broadly, the literature provides a rationale for heterogeneous effects by family structure and baseline constraints. Connelly (1992) shows that higher childcare costs reduce married mothers' labor force participation and emphasizes that selection into childcare arrangements must be treated carefully when estimating these effects. Kimmel (1998) similarly finds that constraints bind particularly strongly for single mothers. In the context of this thesis, these results motivate treating full-day kindergarten as potentially operating through parental time-and-socialization constraints rather than through children's learning alone.

2.3 Parenthood, time costs, and short-run political participation

A second relevant branch of the literature treats political participation as a costly activity. If having young children raises the cost of voting or reduces available time, parental registration rates may fall precisely during the years when households are most constrained. Full-day kindergarten expansions could then increase participation through a symmetric mechanism, by reducing those time constraints among parents of kindergarten-age children.

Cools (2022) provides the most direct evidence for this channel, showing that having an infant reduces voter turnout and interpreting the result as reflecting the time costs of political participation. This result is important because it establishes that family-related time constraints can shift turnout absent any change in political preferences. Teney et al. (2024) complement this by documenting that parenthood shifts voting behavior in ways that are gendered and issue-linked, illustrating how changes in household composition can affect political behavior beyond the pure time-cost mechanism. Finally, Gidengil et al. (2016) provide evidence of intergenerational persistence in political participation, suggesting that changes in parental engagement may have long-run political effects. Although the present analysis targets short-run adult responses, this result helps justify why parental engagement is a meaningful outcome in its own right.

2.4 Visibility, credit-claiming, and the political economy of public goods

The core question of this thesis, whether it is in a politician's electoral interest to expand kindergarten education, is most directly related to the political economy literature on visibility, credit-claiming, and the allocation of public funds. The central argument in this literature is that democratic incentives shape both the level and the composition of public spending, directing resources toward goods that are more visible and more easily attributable to incumbents.

Mani and Mukand (2007) provide the theoretical foundation for this argument, showing that democracy can generate a "visibility effect" that distorts allocation choices toward more observable goods. For kindergarten, this mapping is straightforward: attendance policies are highly salient to the affected families, and delivery is immediate and local, especially in Pennsylvania, where decisions are made at the school-board level. Finan and Mazzocco (2021) develop and test a framework in which electoral incentives shape the allocation of public resources across competing uses, supporting the broader idea that spending choices can be strategically oriented toward electorally productive ends. Gottlieb and Kosec (2019) introduce an important qualification: political competition can improve incentives while simultaneously worsening bargaining efficiency, generating ambiguous predictions for provision levels. The relationship between political competitiveness and kindergarten expansions is therefore an empirical question rather than a theoretical identity.

The literature also warns against over-interpreting electoral responses as clean signals of policy satisfaction. Dynes and Holbein (2020) argue that retrospective voting is "noisy" because voters struggle to map party control onto short-run policy outcomes, so that observed electoral rewards may reflect attribution rather than measured improvements in welfare.

Relatedly, a structural reason why education spending may remain below socially optimal levels even when returns are high concerns the property-tax-based finance system that characterizes most American school districts. Fischel (2001) argues that homeowners treat local public school quality as capitalized into their property values, which can generate both resistance to redistribution and over-investment in highly desirable districts at the expense of poorer ones. Cellini et al. (2010) exploit a regression discontinuity in school bond election outcomes to show that capital investments funded through local referenda do generate positive returns, but that access to this funding mechanism is unequal across districts in ways that correlate with local wealth. These results reinforce the fiscal-sustainability argument developed in Section 10.2: because Pennsylvania school finance relies on local property taxes for more than half of revenues, the capacity to expand full-day kindergarten is structurally unequal across districts.

The gap between social returns to school spending and actual provision is further documented in the broader returns-to-spending literature. Jackson et al. (2016) use variation generated by school finance reforms to identify long-run causal effects of school spending on educational attainment and earnings, finding substantial returns that appear to be systematically underweighted in political spending decisions. Hoxby (2000) examines how competition among public schools affects both quality and taxpayer costs, finding that competitive pressure can improve efficiency but does not by itself resolve the underlying funding inequalities that constrain provision in low-wealth districts. Bursztyn (2016) brings these threads together in the Brazilian context, showing that politically-induced misalignment between provision and social returns persists even when the returns are objectively large, reinforcing the general lesson that the political economy of education is not self-correcting.

2.5 Electoral incentives and the adoption of early childhood programs

A final, more targeted literature examines early childhood program adoption explicitly through the lens of electoral incentives and institutional rules. This strand is most directly relevant to the first part of this thesis, which studies correlates and potential determinants of adoption.

Tebaldi, Gassmann, et al. (2024) use a regression discontinuity design in Brazil to estimate the effect of re-election incentives on the adoption of an early childhood development program, providing direct causal evidence that politicians' incentives shape early childhood policy. Tebaldi and Portella (2025) exploit an electoral rule threshold to study how institutional structure affects early childhood education provision and spending. Finally, Marchese et al. (2025) show that external shocks can alter the salience of childcare spending and that political actors adjust provision accordingly.

The predictors of full-day kindergarten adoption have also been studied directly. Greater Democratic legislative control is positively associated with FDK enactment (Curran, 2015), as is a higher share of women in the legislature (Cohen-Vogel et al., 2020). Among economic predictors, higher local unemployment and lower district income are both associated with earlier adoption (Cohen-Vogel et al., 2020; Yang, 2023). Yang (2023) further finds that state funding typically constitutes a large share of FDK financing, and that smaller, lower-achieving districts are more likely to adopt. Taken together, these findings suggest that economic and fiscal factors are stronger predictors of adoption timing than partisan composition alone, a conclusion consistent with the endogeneity assessment in Section 6.2 of this thesis.

3. Institutional context and policy variation

3.1 U.S. kindergarten policy landscape

In the United States of America, decisions concerning educational policies are delegated to each individual state. As a result there is no common policy, which generates substantial variation across the country. According to the Education Commission of the States, 17 states and the District of Columbia mandate that children attend kindergarten prior to entering first grade. Additionally, “16 states and the District of Columbia require districts to offer full-day kindergarten, and 44 states plus the District of Columbia require districts to offer at least half-day kindergarten." (Fischer et al. (2023)). However, only two states, Louisiana and West Virginia, mandate full-day kindergarten attendance. 6 other states delegate kindergarten entrance age, provision, and attendance policies to local educational agencies: Colorado, Massachusetts, New Hampshire, New Jersey, New York, and Pennsylvania.

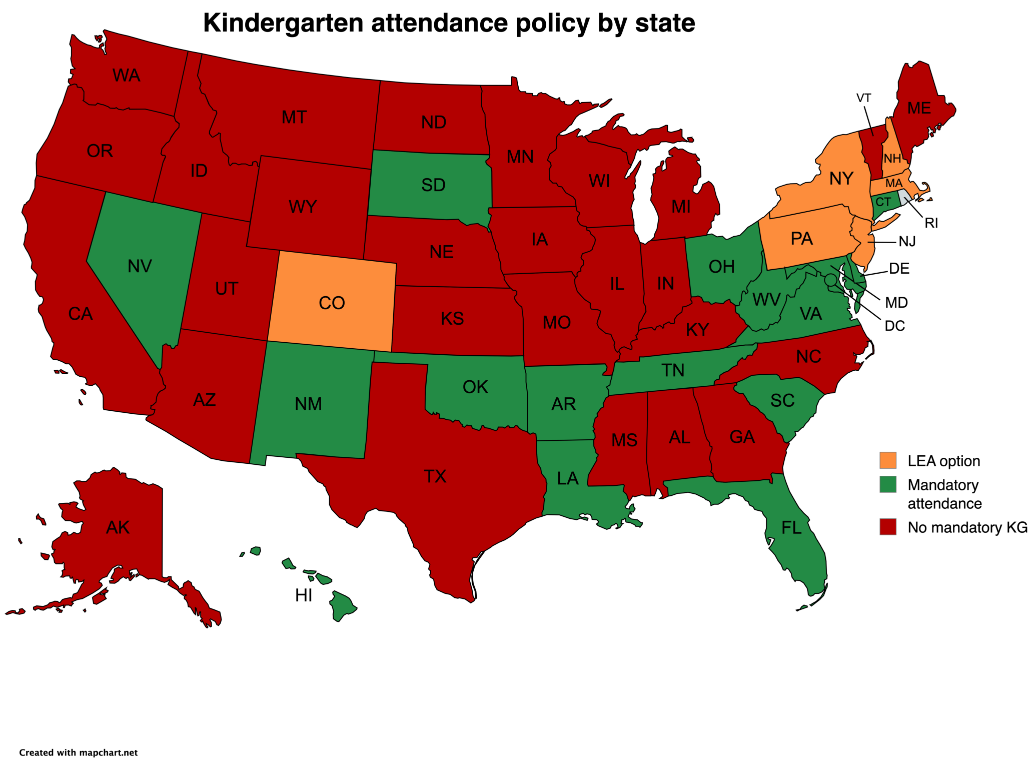

Figure 3.1: Attendance policy by state

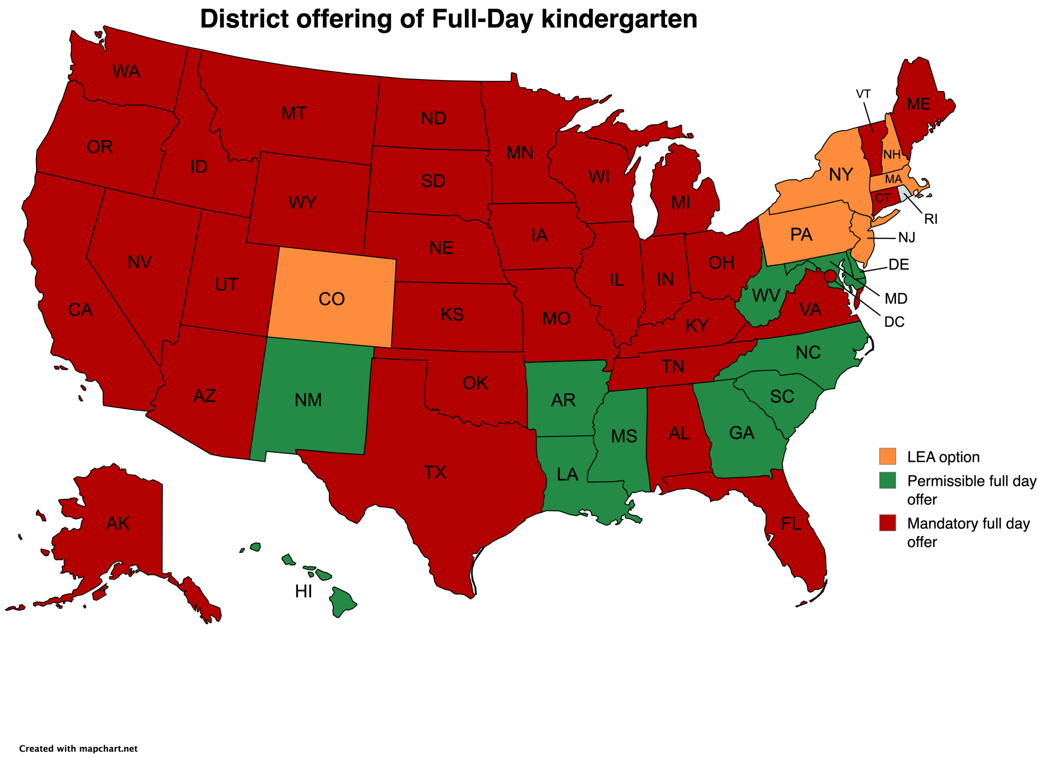

Figure 3.2: Full-day offer by state

Based on where they reside, and on what parents want for their children, when a child reaches kindergarten age (most often 5 years old) they have three possible options. The first, if kindergarten attendance is not mandatory, is to not attend any form of preschool. This implies the parents either having to take care of their children themselves all day, or finding some form of childcare. The second option consists of enrolling children in half-day kindergarten programs, where available. The duration of such programs varies across states, but typically ranges between two and four hours of daily instruction. The academic content covered generally mirrors that of full-day programs, albeit in a more condensed format. Half-day kindergarten is broadly regarded as a suitable alternative for families seeking to allow their children an additional transitional year before committing to a full-day academic schedule. It should be noted, however, that this arrangement necessarily requires parents or guardians to provide care for the remainder of the day. The third and final option is full-day kindergarten, which typically constitutes a six-to-seven-hour daily program. This format affords children greater opportunities for early socialization and allows academic content to be delivered at a more measured pace, while simultaneously offering families a full-time childcare solution.

3.2 Pennsylvania: LEAs and kindergarten provision

Pennsylvania is one of six states which fully delegates kindergarten policies to local educational agencies. Pennsylvania has 500 school districts and 667 local educational agencies. In practice, given the composition of school boards, decisions are more often taken at the school district level. In total, over 1.7 million children per year are enrolled across grade levels, spanning from kindergarten through to the final year of secondary education, 12th grade. School districts (SDs) vary considerably in size, with the smallest (Austin Area) having 162 students enrolled in total in 2024, including only 7 in kindergarten. The largest, Philadelphia, had 117,985 students enrolled in 2024, including over 8,000 in full-day kindergarten. There are on average 3,031 students per school district (median of 1931).

Pennsylvania schools obtain their funding from a combination of federal, state, and local government sources. The federal share of funding was of roughly 13.6% in the 2021-2022 school year. The state itself covers 35.1% of the expenses, and the remaining 51.3% comes from local taxes, such as property taxes (source: National Center for Education Statistics). The total funding amounts to roughly $38.2 billion, and due to the large local proportion of funding there are large disparities across the state, as shown in the map below. This thus entails a differing ability to provide certain services such as universal full-day kindergarten. It is thus important to account for this endogeneity when performing future analyses.

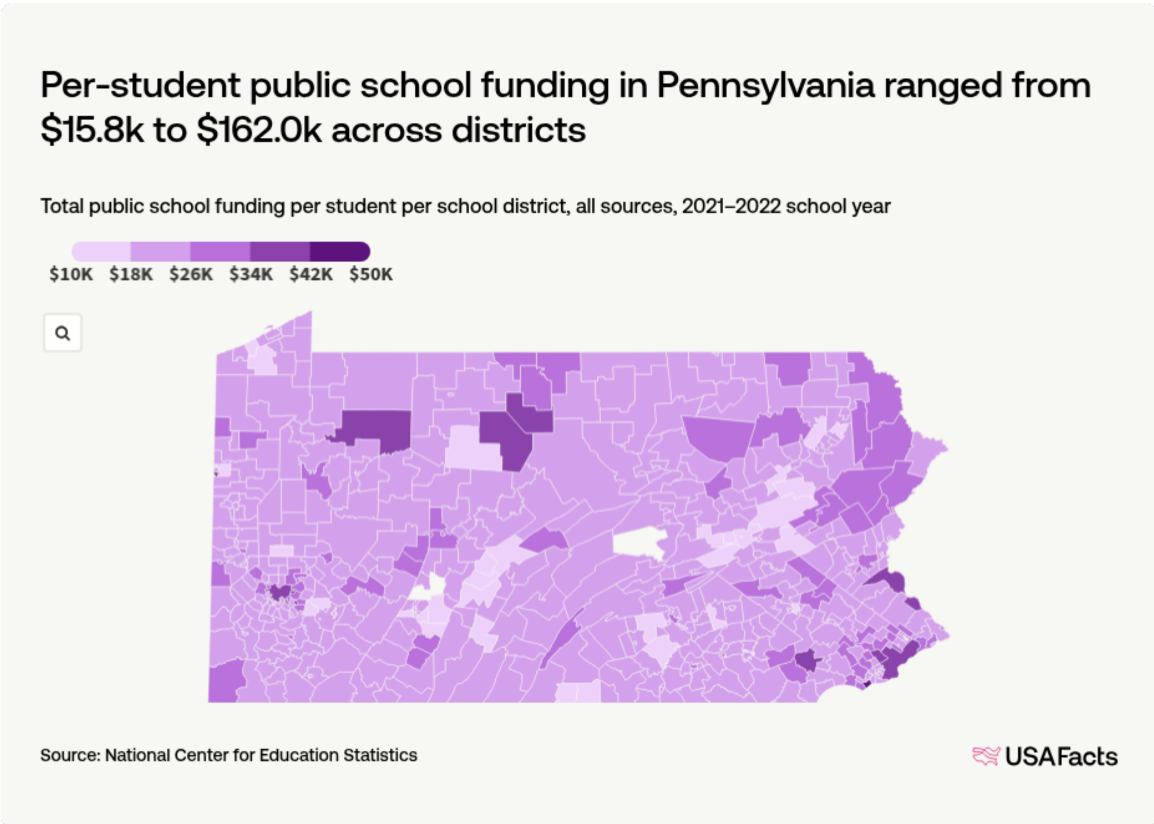

Figure 3.3: School funding in Pennsylvania

3.3 Relevant political actors

A few different elections and political actors are important in the Pennsylvanian context. They are responsible to some extent for setting the policies/funding or other related factors.

The first key actors are the members of the local school boards. Their function is to exercise direct budgetary sovereignty over district expenditures. Because Pennsylvanian law does not mandate kindergarten and leaves it up to the LEAs, school boards treat it as a discretionary expense. Boards are responsible for deciding both the “quantity" (half-day vs. full-day) and the “quality" (student-teacher ratios for instance) of the service provided. The elections happen in odd years.

The second key actors are the members of the PA General Assembly, that is the state senate and the state house. Their responsibility in this domain is to determine the Basic Education Funding (BEF) Formula and the state subsidies for kindergarten related projects. Since the legislators control the redistributive mechanism, they effectively decide how much state tax revenue is transferred to "high-need, low-wealth" districts to offset local tax base deficiencies. The elections that happen in even years can thus also reflect the satisfaction of the constituents with the policies in place, including school finance ones.

Finally, the last key actor is the Governor, who is responsible for setting the Fiscal Agenda and executive priorities. The Governor's budget proposal can be seen as the opening bid for early childhood investment, and depending on the priorities of the office can try to push for a certain legislative agenda.

This paper will primarily focus on voter registration, controlling for incumbency in the legislative districts. It will also examine the channels through which potential change may occur, specifically changes in labor force participation. In terms of electoral result, a mapping will be made between the school districts and the relevant state Senate and state House ones, in order to study the potential political dividend of such a reform

4. Data sources

4.1 Units of observation and key variables

This paper consists of two main analyses. The first is a descriptive analysis over time of changes in intensity of kindergarten adoption in the USA, looking at information at the state level, from the mid 80s to today. In the second, causal part the treatment considered, that is adoption of mandatory full-day kindergarten, is defined at the district level. The analyses will be done at that level in order to study the impact of changes in kindergarten policy on the probability of registering to vote, for people of age to have kindergarten-age children. These analyses will be aggregated at the district rather than the individual level in order to be able to use population wide denominators and override the issue of the absence in the data of people who never register to vote.

Treatment is defined as a discrete and persistent switch from a district having <70% full-day enrollment to >95% the following year and every year thereafter. The switch must thus occur instantly, and districts which may progressively implement the policy are not considered, due to the less clean treatment. Additional groups are also defined, with the always-treated districts corresponding to those which always have more than 95% enrolled in FDK, and the never-treated those who always have enrollment levels under 50%. In the data we have 277 Always-Treated districts, 142 treated, 53 Never-Treated, and 28 unassigned. Treatment is recovered through the data, with a series of validity checks also being performed (checking district by district for communication around the imputed date concerning the expansion when available). See section 5.2.3 for a more detailed analysis on the validity checks.

The main outcomes studied are political. The first one is whether people are more likely to register to vote following the change in policy. This consists of looking at people aged 25 to 44 (the estimated bounds of the age of parents of 5 year olds) and seeing the share that register in each year in a district, and observing possible discontinuities. This analysis is then performed by looking at the share for each party of the enrollments in a given year, compared to the switch year.

Additional analyses look at these changes conditioning on incumbent status. Furthermore, the electoral effects on the composition of the state legislature are also presented, specifically in terms of vote margins and wins in the state House and state Senate races.

4.2 Data sources and construction

The main data for voter information comes from the Pennsylvania board of elections. It contains information including voter IDs, date of birth, exact address, all the elections in which an individual can vote (including the school district ones which facilitates the mapping), and the exact date at which a person registered to vote. It contains 8911681 observations, covering every individual currently registered to vote in Pennsylvania. It also contains the party registration history, as well as the participation or lack thereof in the most recent elections (such as primaries, midterms, and presidential elections).

Information on kindergarten attendance by type and LEA, and enrollment per grade level comes from the PA department of education. For the same information at the state level for every state, data from 2010 to 2025 comes from the American Community survey (ACS). Population data from 1980 to 2010 comes from archives of the census, and kindergarten enrollment information comes from U.S. Department of Education, National Center for Education Statistics, ELSi generator. Political results for legislative composition at the state level from 2009 to 2025 comes from the National Conference of State Legislatures, and state level controls such as median income and unemployment rates also come from the ACS/Census. More precise political estimates, including for each race within the state comes from the KlarnerPolitics database "State Legislative Election Returns, 1967-2016". This gives the winner of and number of electors in all of the electoral contests within Pennsylvania.

The data used to obtain the possible predictors comes from the Current Population survey, as does the labor force participation rates for women both married and not.

4.3 Descriptive statistics

A simple analysis is run to compare the districts that are never-treated, always-treated, and treated early or late. The characteristics of these are presented in the balance table below

	Early switcher	Late switcher	Never-treated	Always-treated
Registration				
N	54	40	47	232
Pre-treatment reg. rate	0.0303	0.0458	0.0585	0.0263
Demographics				
Population 25–44 (mean)	6726	6908	8671	4548
Female share	0.510	0.513	0.512	0.507
Partisan composition (ages 25–44)				
Democrat share	0.341	0.332	0.359	0.307
Republican share	0.487	0.478	0.455	0.543
Fiscal				
Median household income ($)	57732	67218	89269	47921

Table 4.1: Pre-treatment balance by district group. Treated districts split at the median treatment year. Registration rate = new registrations / ACS population 25–44 (B01001). Party shares and female share computed from the voter file among adults aged 25–44 in each district's pre-treatment period. Median household income from ACS B19013, averaged over pre-treatment years. Always-treated baseline uses 2010–2014 as a proxy pre-period. Only districts included in the ACS data are present in this table, leading to different N than the overall data, and potential representativity issues.

This table shows that the always-treated districts are slightly smaller, and the never-treated slightly bigger than the treated districts, which is consistent with the idea that it is easier to provide daycare for children if there are fewer of them. However, the data used to generate this table comes from the ACS, which does not contain every single school district (only the big enough ones), so there may be a slight selection bias among those that are actually presented here. The descriptive statistics must thus be interpreted cautiously. Interestingly enough the always-treated districts have the lowest median income, which may seem somewhat counterintuitive given the idea that schools are largely financed through property taxes. However, as the Basic Education Funding formula used in Pennsylvania provides subsidies for families living under up to 184% of the federal poverty line, which is set at $33,000 in 2026 for a family of 4, it seems plausible that since the median income is under that threshold (47,921<60,300), the schools obtained state level subsidies, enabling this earlier treatment. The never-treated on the other hand may have a wider available supply of private child care, which explains why they would perhaps not implement such a policy. In light of the possible differences between the districts though, it is especially important to ensure that the parallel trends assumption (which will be developed further in the next section) holds, for there to be some comparability.

5. Empirical strategy and identification

5.1 Part 1: Descriptive understanding of politics in kindergarten policies

The first part of this thesis is explicitly descriptive and should be interpreted as such. Rather than attempting to establish a causal link between kindergarten expansion and political outcomes at this stage, the goal is twofold: first, to document the evolution of full-day kindergarten enrollment across the United States from the mid-1980s to the present, and secondly to assess whether the timing of state-level adoption of mandatory attendance or full-day offer laws is predicted by pre-existing political variables. This second objective is not merely descriptive, it directly informs the identification strategy developed in Part 2. The covariates considered are those suggested by the literature as potentially having a predictive impact. As not all states have state wide policies, the sample size of switching states is somewhat limited.

5.1.1 Exposure intensity and scaling

The unit of observation in Part 1 is the state-year. The key variable of interest is the share of kindergarten-age children enrolled in kindergarten within a given state and year, which is used as a continuous measure of FDK exposure intensity. At the state level, especially for the older dates, the data does not allow for the distinction between full and half day attendance, so the results found are more of an exploration of the political effects of attendance than of FDK. Formally, let

$$\text{pct\_in\_kg}_{st} = \frac{\text{KG Enrollment}_{st}}{\text{KG-age Population}_{st}}$$

where kindergarten enrollment comes from the NCES ELSi generator (1986–2010) and the American Community Survey (2010–2024), and the denominator is the estimated number of 5 year olds in state s in year t. Since, apart for the data from 2000 to 2009 which contains population information for every single age, age-specific population data from the Census and ACS are typically reported in 5-year bins rather than single years, the 5-to-9 age bin is divided by 5 to recover an approximation of the count for kindergarten-age children. This approximation is standard in the literature and is assumed to introduce minimal error given the smoothness of age distributions within this range. A three-year rolling average of this intensity measure, $\overline{\text{pct\_in\_kg}}_{st}$, is constructed as the primary exposure variable for the two way fixed effect regressions in Section 5.1.2. This smoothing reduces the influence of year-to-year measurement noise in enrollment data. It also captures the accumulated rather than instantaneous effect of FDK availability on political outcomes. Both the simple nationwide average and the population-weighted average (obtained by weighting each state by its estimated kindergarten-age population, in order to avoid overrepresentation of small states) are computed and reported to verify that the trends are not driven by large states.

5.1.2 Estimates of correlates

To examine whether kindergarten enrollment intensity correlates with partisan political outcomes at the state level, a two-way fixed effects model is estimated. The baseline specification is:

$$Y_{st} = \beta \cdot \overline{\text{pct\_in\_kg}}_{st} + \alpha_s + \gamma_t + \varepsilon_{st}$$

where Y_{st} is a political outcome for state s in year t, α_s is a state fixed effect which absorbs all the time-invariant differences across states, and γ_t is a year fixed effect which absorbs the time trends which are common to all states. The coefficient of interest β therefore identifies the within-state, within-year variation in kindergarten intensity as a predictor of political outcomes. Standard errors are clustered at the state level to account for serial correlation within states over time.

Three types of political outcome variables are considered. The first consists of measures of partisan control of the state legislature: the Democratic seat share in the state Senate ($\text{lean\_senate}_{st} = (\text{Dem seats} - \text{Rep seats})/\text{Total seats}$) and in the state House. These are weighted by the total number of seats in each chamber to account for differences in legislature size across states. The second family concerns the partisan affiliation of the governor, measured as a binary indicator equal to one if the governor is a Democrat. The third is a composite trifecta indicator, measuring full Democratic or Republican control of both chambers and the governorship. For the legislative seat share outcomes, specifications are run both in levels (net Democratic seat advantage) and as a share (Democratic lean), and three definitions of the exposure variable are tested: the contemporaneous intensity, the three-year rolling average, and the cumulative exposure since 1986. The variables are coded as pro-Democrat outcomes being equal to 1 and pro-Republican being equal to 0 as that is the direction suggested by the literature, but of course the reverse coding would also be possible.

Importantly this TWFE specification is purely correlational. It does not aim to establish a causal link between kindergarten attendance and partisan outcomes at the national level. The identification in this part relies on within-state variation in kindergarten attendance intensity over time, but this variation is not plausibly exogenous: states may have expanded kindergarten precisely because they had Democratic governments, or because economic conditions simultaneously increased both demand for kindergarten and Democratic (or Republican) vote shares. The results from this section should therefore be interpreted as suggestive evidence of an association, which is contextualized and qualified by the endogeneity assessment below.

To establish the degree of endogeneity in adoption timing, two additional analyses are performed. First, pre-adoption enrollment trends are plotted separately by the partisan composition of the state legislature in the years leading up to mandatory attendance law adoption. If adoption is politically driven, Democratic-controlled legislatures should have steeper pre-adoption enrollment growth than Republican-controlled ones. Second, a logit model is estimated where the outcome is an indicator for having adopted a mandatory full-day kindergarten law by a given year, and the predictors are the pre-period Democratic legislature share, a Democratic governor indicator, and the pre-period state-level female labor force participation rate (standardized). Formally:

$$\Pr(\text{Adopted}_s) = \Lambda\left(\delta_1 \cdot \text{DemLegis}_{s,\text{pre}} + \delta_2 \cdot \text{DemGov}_{s,\text{pre}} + \delta_3 \cdot \overline{\text{LFP}}_{s,\text{pre}}\right)$$

where $\Lambda(\cdot)$ denotes the logistic CDF and pre-period averages are computed over the years prior to 2000 for states that have not yet adopted. Insignificant coefficients on the political variables would support the claim that, conditional on economic characteristics, the timing of adoption is not strongly determined by partisan control, a precondition for the causal analysis in the following part.

5.2 Part 2: Rigorous identification, case study of Pennsylvania

5.2.1 Conceptual identification problem

This section, starting from the descriptive analysis conducted in the previous section aims to reestablish the causal link. In order to do so, it focuses on the specific case study of Pennsylvania, which was previously motivated in section 3. The central empirical challenge of this thesis is that the adoption of full-day kindergarten by a school district is not randomly assigned. Districts that choose to expand FDK may differ systematically from those that do not, in ways that are also likely to predict the political outcomes of interest. Three distinct sources of confounding are worth making explicit.

The first is selection on levels. Districts that expanded FDK earliest in Pennsylvania tend to be smaller and poorer, as suggested by the balance table reported in Section 4.3. These characteristics are also associated with lower baseline voter registration rates, lower baseline Democratic registration. A naive comparison of registration rates between FDK and non-FDK districts could therefore conflate the effect of FDK with these pre-existing differences. The second concern is selection on trends. Even after conditioning on levels via fixed effects, districts could be on different trajectories. If civic engagement and FDK adoption are jointly driven by a common underlying trend, then even a within-district comparison would overstate the causal effect of FDK. The third concern is strategic timing. School boards, which are the relevant decision-making bodies in Pennsylvania, hold elections in odd-numbered years. If boards adopt FDK strategically in anticipation of electoral cycles, for instance, expanding the program just before an election to court parents of young children, then the timing of adoption is endogenous to the very electoral outcomes this thesis seeks to explain. In that case, any estimated positive effect on voter registration could reflect reverse causality.

These concerns jointly imply that naive comparisons, whether cross-sectional, first-difference, or simple event-time plots, are likely to produce misleading estimates. A credible identification strategy must find variation in FDK adoption that is plausibly orthogonal to the potential outcomes. The staggered adoption of FDK across Pennsylvania school districts, combined with the modern difference-in-differences framework described below, provides such variation.

5.2.2 Modern staggered DiD / event-study framework

Recent methodological advances in the difference-in-differences literature have demonstrated that the standard two-way fixed effects (TWFE) estimator is biased when treatment effects are heterogeneous across cohorts or over time (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Chaisemartin and D'Haultfœuille, 2020; Borusyak et al., 2024). The intuition is straightforward: in a staggered adoption setting, TWFE implicitly uses already-treated units as controls for not-yet-treated units. If the treatment effect grows over time or differs across adoption cohorts, this produces contaminated comparison groups and can even reverse the sign of the estimated effect.

This thesis therefore uses the estimator proposed by Callaway and Sant'Anna (2021) (CS DiD) as its primary identification strategy (although some of the other estimators will later serve for robustness checks). This strategy computes cohort-specific treatment effects and aggregates them in a second step. The framework is built around the following potential outcomes notation. Let $Y_{it}(g)$ denote the potential outcome for unit i at time t if that unit's treatment group is g, where $g \in \mathcal{G} \cup \{0\}$ and $g = 0$ denotes the never-treated group. The observed outcome is $Y_{it} = Y_{it}(G_i)$, where G_i is the year in which unit i first receives treatment (its adoption cohort), with $G_i = 0$ for never-treated units. The treatment indicator is $D_{it} = \mathbf{1}[t \geq G_i, G_i > 0]$.

The parameter of interest at the cohort-by-time level is the group-time average treatment effect:

$$\text{ATT}(g, t) = \mathbb{E}\left[Y_{it}(g) - Y_{it}(0) \mid G_i = g\right]$$

which measures the average effect on units first treated in year g, evaluated at calendar time t. Note that $\text{ATT}(g, t)$ is well-defined for both pre-treatment periods ($t < g$), where it should be zero under parallel trends, and post-treatment periods ($t \geq g$), where it captures the dynamic effect of FDK adoption.

The key identifying assumption is unconditional parallel trends: the average potential outcome under no treatment would have followed the same trend in treated and untreated units. Formally:

$$\mathbb{E}[Y_{it}(0) - Y_{it-1}(0) \mid G_i = g] = \mathbb{E}[Y_{it}(0) - Y_{it-1}(0) \mid G_i \in C]$$

for all t and all comparison groups C. The primary comparison group throughout is the set of not-yet-treated units, that is, districts that have not yet undergone the FDK transition by period t but will eventually do so. This group is preferred over the never-treated districts as the control, because with only 53 never-treated districts against 142 treated districts, the never-treated constitute a smaller and potentially less comparable set. Districts that never expanded FDK at all may differ structurally from those that eventually adopted, in ways that are difficult to fully account for even within a doubly-robust framework. The not-yet-treated comparison group draws on the much larger pool of eventual adopters observed prior to their own transition, making the parallel trends assumption more plausible and the weighting more stable. The never-treated group is retained as a robustness check, and results are reported for both control group definitions in Section 9.2.3.

The $\text{ATT}(g, t)$ parameters are estimated via the doubly-robust (DR) estimator of Callaway and Sant'Anna (2021), which combines inverse probability weighting (IPW) on the propensity score with a regression adjustment on the outcome model. This estimator is consistent if either the propensity score model or the outcome regression is correctly specified, providing additional insurance against misspecification of either component.

5.2.3 Treatment definition

A school district is classified as treated in year g if it transitions from fewer than 70% of kindergartners enrolled in FDK in year $g - 1$ to more than 95% in year g, and maintains that level in all subsequent years. This two-threshold rule identifies discrete, persistent switches rather than temporary fluctuations in enrollment.

Three distinct groups are defined in the data. The 142 *treated* districts are those that satisfy the threshold rule at some point during the sample period and constitute the treated population. The 53 *never-treated* districts are those that never exceeded 50% FDK enrollment throughout the panel, and serve as the secondary control group in robustness specifications. The 277 *always-treated* districts are those that already exceeded 95% FDK enrollment at the start of the observation window and are excluded from the DiD estimation, since their pre-treatment period cannot be observed. A residual group of 28 districts could not be unambiguously classified and are also excluded.

Treatment is recovered from administrative enrollment data rather than from official policy records. This approach is motivated by the fact that school board decisions in Pennsylvania are not centrally archived. A series of validity checks was performed: districts with large year-on-year enrollment changes (Δpct_in_kg $>$ 0.20) were flagged and manually cross-referenced against available school board communications and local press archives to verify that the imputed treatment year corresponds to an actual policy shift, when available. As many of the districts switched in early years or are relatively small, there do not necessarily exist online traces of communication surrounding the treatment for all of them. For the 57 treated districts for which there was communication, the imputed treat date was verified. Half of the communication was in local press (29 of 57 districts), 17 came from district communication, 7 from Facebook threads, and the rest come from other sources (LinkedIn, political programs, etc.). Treatment classification is robust to alternative threshold specifications of a 50% jump, as reported in Section 9.2.3.

5.2.4 Baseline specification

The primary outcome in Part 2 is the new voter registration rate among adults of plausible parenting age. Specifically, for district i in year t, the outcome is:

$$\text{reg\_rate}_{it} = \frac{\text{New registrations}_{it}^{[25,44]}}{\text{Pop}_{it}^{[25,44]}}$$

where the numerator is the number of individuals aged 25–44 who registered to vote for the first time in a given district and calendar year, taken from the Pennsylvania Board of Elections voter file, and the denominator is the ACS 5-year estimate of the population aged 25–44 in that district (constructed from B01001, summing four exact 5-year bins: 25–29, 30–34, 35–39, 40–44). The 25–44 age window was chosen as this range corresponds to the plausible ages of parents of 5-year-old children, assuming that most people give birth between the ages of 20 and 39. As the voter file only includes people who actually registered to vote at some point before December 2025 (when the data was downloaded), a gap remains for the people who never register, hence why it is important to take the ACS as a population denominator instead of the cell size.

The cohort-specific ATT estimates are aggregated into a single scalar via:

$$\text{ATT}^{\text{simple}} = \sum_{g \in \mathcal{G}} \sum_{t \geq g} w_{gt} \cdot \text{ATT}(g, t)$$

where w_{gt} is proportional to the cohort size (weighted by the ACS population aged 25–44), ensuring that larger districts contribute more to the aggregate.

5.2.5 Event-study and dynamic effects

To examine the dynamic pattern of treatment effects and to assess the plausibility of the parallel trends assumption in the pre-treatment period, the cohort-time estimates $\text{ATT}(g, t)$ are also aggregated into a single event-study path indexed by relative time $e = t - g$:

$$\text{ATT}^{\text{es}}(e) = \sum_{g \in \mathcal{G}} \mathbf{1}[g + e \in \mathcal{T}] \cdot w_g \cdot \text{ATT}(g, g + e)$$

This is implemented with a balanced panel restriction: only cohort-time cells for which all cohorts contribute an observation are included, ensuring that the event-study path does not change composition as the horizon extends. The pre-treatment period is binned at $e = -6$ and the post-treatment period is balanced at $e = +10$. Binning the pre-period prevents the point estimate at $e = -6$ from absorbing variation from earlier periods through the reference category normalization.

The event-study path serves two purposes. First, the pre-treatment coefficients at $e \in \{-6, -5, \ldots, -2\}$ are a direct test of parallel pre-trends: under the null of parallel trends and no anticipation, these estimates should be jointly zero. Second, the post-treatment path at $e \in \{0, 1, \ldots, +10\}$ documents whether effects emerge immediately, build over time, or fade, which is informative for the interpretation of the mechanism.

Two sets of event-study estimates are produced: for the pooled sample (all adults 25–44), and separately for female (mothers) and male (fathers) adults, using gender-specific population denominators from the ACS. The gender decomposition is central to the mechanism analysis, since the FDK as childcare hypothesis predicts that effects should be strongest for mothers, for whom the childcare constraint is typically more binding.

The main channel which will be explored is changes in labor force participation for women by marital status, as single mothers are the ones expected to gain the most. For the labor force participation analysis in Part 2b, the same CS DiD framework is applied at the district-year level, using ACS 5-year estimates as both outcome and denominator. The outcome is the maternal LFP rate conditional on family structure:

$$\text{LFP rate}_{it}^{(m,k)} = \frac{\text{Mothers in LF, family type } m, \text{ child age } k}{\text{Children in arrangement } m, \text{ age } k}$$

estimated separately for single mothers (m = single) and two-parent families (m = two-parent), and for children under 6 (k = under 6, the treatment cell) and children aged 6–17 (k = 6-17, the placebo). Since the data does not separately identify the employed and unemployed mothers, the outcome is a labor force participation rate, not an employment rate. The control group for the LFP analysis is likewise the not-yet-treated group, consistent with the main registration specification.

Electoral outcomes are explored using the same Callaway-Sant'Anna specification for Pennsylvania House districts, where the sample is large enough to support the staggered DiD design. For Pennsylvania Senate districts, the smaller number of units makes the CS DiD approach impractical. Instead, a two-stage imputation estimator in the spirit of Borusyak et al. (2024) is applied to the continuous intensity setting. In the first stage, unit and time fixed effects are only estimated on untreated observations, to recover an imputed counterfactual $\hat{Y}_{dt}(0)$ for all district-election cells. In the second stage, the residual $Y_{dt} - \hat{Y}_{dt}(0)$ is regressed on the interaction $\text{Intensity}_d \times \text{Post}_{dt}$, where $\text{Intensity}_d \in [0, 1]$ is the time-invariant share of the Senate district's registered voters residing in treated school districts, and Post_{dt} is an indicator for election cycles at or after the district's first treatment year. The coefficient on this interaction is interpreted as the effect of moving from zero to full FDK exposure within a Senate district. An event-study version is also estimated, replacing the single post indicator with a set of relative-election-cycle dummies interacted with Intensity_d, with $t = -1$ omitted as the reference period. Standard errors are clustered at the district level throughout.

5.3 Inference

Most of the CS DiD estimates are accompanied by simultaneous confidence bands constructed via the multiplier bootstrap with 999 iterations, as described in Callaway and Sant'Anna (2021). Simultaneous confidence bands differ from the more typical pointwise bands in an important respect: they control the probability that the entire event-study path falls within the band, rather than controlling coverage at each period independently. This is the appropriate measure of uncertainty for event-study plots, where one is jointly assessing the pre-trends and the post-treatment path. Pointwise bands (based on $1.96 \times \text{SE}$ at each period) are also reported for transparency and robustness, since they allow comparison with other estimators that do not produce bootstrapped critical values.

Standard errors are clustered at the district level in all specifications. This accounts for serial correlation in the outcomes within districts over time, which is particularly important given that the voter file panel spans up to 15 years. For the state-level ITT regressions in Part 1, standard errors are clustered at the state level.

To assess robustness of the inference to the choice of estimator, two alternative approaches are benchmarked against the primary CS DiD. The first is the stacked DiD estimator (Baker et al., 2022): for each adoption cohort g, a sub-panel is constructed containing the cohort itself and all districts that were not yet treated by $g + K$ (where K is the maximum post-treatment horizon), and these sub-panels are stacked. A standard TWFE regression with unit-by-stack and period-by-stack fixed effects is then estimated on the stacked dataset. This approach avoids negative-weighting bias without requiring semiparametric assumptions, at the cost of duplicating not-yet-treated observations across stacks. The second is the two-stage imputation estimator of Borusyak et al. (2024) (BJS): in the first stage, unit and time fixed effects are estimated on the untreated observations only (pre-treatment periods for treated districts and all periods for not-yet-treated districts at the relevant horizon), producing a predicted $\hat{Y}_{it}(0)$ for all observations. In the second stage, the residual $Y_{it} - \hat{Y}_{it}(0)$ is regressed on relative-time indicators for the treated units, with clustered sandwich standard errors. Agreement across these three estimators is the primary robustness criterion for the main voter registration results.

6. Analysis and findings I: State-level adoption patterns (descriptive ITT)

This chapter presents the first of the two empirical analyses. As noted in the empirical strategy, all results in this section are explicitly descriptive and should be interpreted with caution. Given that states can self-select into adopting mandatory kindergarten, the associations documented here cannot be interpreted as causal. The goal of this chapter is therefore twofold. The first is to document how kindergarten enrollment has evolved across the United States over a period of roughly four decades, and to characterize which states adopted mandatory attendance laws and when. The second, and perhaps more important goal from the perspective of the identification strategy developed in Chapter 5, is to assess whether that adoption timing is systematically predicted by preexisting political characteristics. As will be discussed, this assessment has direct implications for the credibility of the causal analysis in Chapter 7.

6.1 Adoption timing and mapping

Figure 3.1 in Chapter 3 already documented the cross-sectional picture as of the most recent data: 17 states and the District of Columbia mandate kindergarten attendance, a further 16 mandate a full-day offer without requiring attendance, and the remaining states delegate these decisions to local educational agencies or impose no requirement at all. What this static picture does not convey, however, is the extent to which the national landscape has changed over time, and how rapidly kindergarten has become the norm even in states that have not passed formal mandates.

Figure 6.1 plots the event study of kindergarten enrollment around the year a state enacted a mandatory full-day kindergarten offer law ($t = 0$), restricting to states that adopted after 1986. Two series are shown: the simple nationwide average of the enrollment share and a version weighted by each state's estimated kindergarten-age population. Importantly, this does not necessarily mean that all children were in full-day kindergarten, just that the districts had to give the families the options. In these states, provided both are offered, the parents can thus choose the education/childcare offer which best fits their needs. The profile of the curves tells an interesting story. The pre-adoption trend is increasing, but the post-adoption curve seems steeper. This suggests that even just providing the offer for full-day, without mandating it, leads to an increase in enrolled children. We see that the states that switch to a mandatory FDK offer have an attendance which seems to reach 100% after roughly 12 years. [1]

[1] Due to the different origins of the data and the necessary imputation of kindergarten age children, compared to the actual number enrolled which is known, there are occasional slight discrepancies in the data, explaining the few values over 100%.

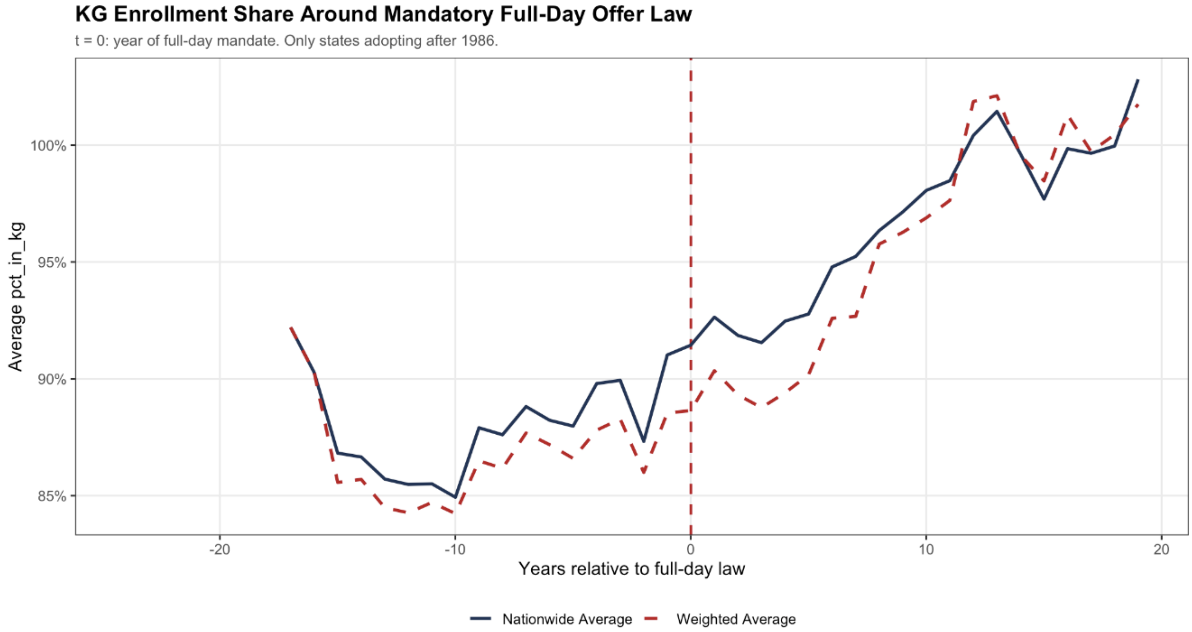

Figure 6.1: KG enrollment share around mandatory full-day offer law enactment

6.2 Predictors of adoption and potential endogeneity

It is next important to study whether the timing of mandatory FDK adoption is politically driven. This is an interesting descriptive question, but also important to consider, as it could be a direct threat to the causal analysis in Chapter 7. If states adopted FDK when and because they had Democratic governments, then any estimated association between FDK intensity and Democratic electoral outcomes could simply reflect reverse causality: Democratic governments both expand FDK and win elections, for reasons unrelated to the policy itself.

To assess this threat, Figure 6.2 plots pre-adoption FDK enrollment trends separately for states with Democratic, Republican, and split legislative control in the years leading up to mandatory attendance law adoption. If adoption were politically driven in a systematic way, one would expect to see divergent trends across these three groups: Democratic-controlled legislatures should have higher and faster-growing enrollment even before formally mandating attendance. The figure suggests that this is not the case. The three series track each other relatively closely throughout the pre-period, with no meaningful divergence in either level or slope. This near-parallel pre-trend across partisan groups is the first piece of evidence that adoption timing is not strongly politically determined (at least in terms of raw data, this does not add in controls).

Figure 6.3 formalizes this with a logit regression. The outcome is an indicator for whether a state had adopted a mandatory full-day kindergarten law, and the predictors are the pre-period Democratic legislature indicator, a Democratic governor indicator, and the pre-period state-level female labor force participation (LFP) rate, standardized to have mean zero and unit standard deviation. The female LFP measure is drawn from the IPUMS CPS ASEC and averaged over years prior to 2000 for states that had not yet adopted by that point. It is included as an economic predictor motivated by the literature discussed in Chapter 2: states with higher baseline female LFP may face stronger demand for publicly

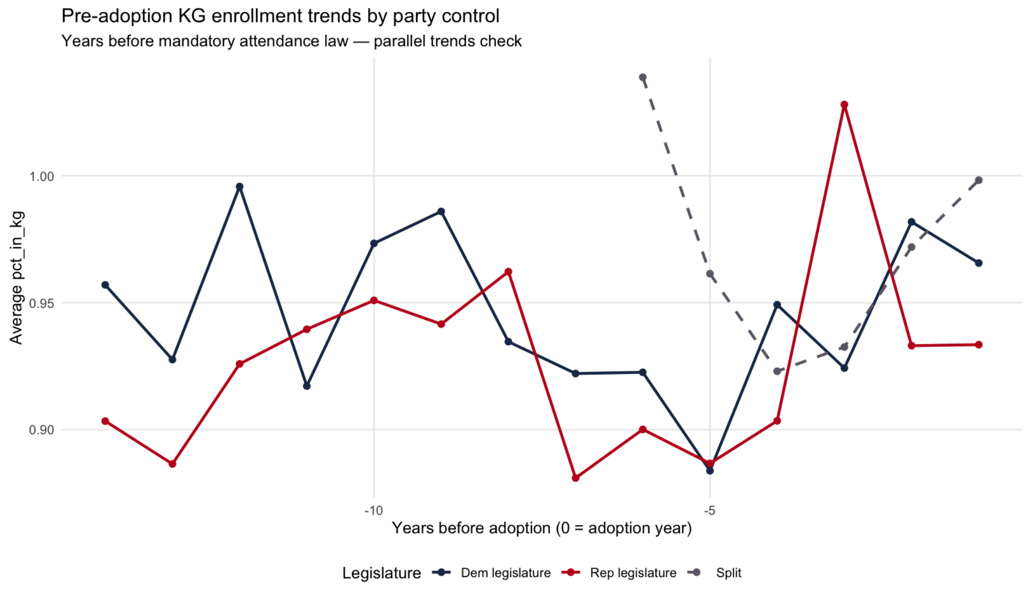

Figure 6.2: Pre-adoption KG enrollment trends by legislative party control

provided childcare, and may therefore be more likely to expand FDK regardless of partisan composition. Low female LFP may also be indicative of a poor pre reform childcare offer, therefore the sign will be interesting to consider.

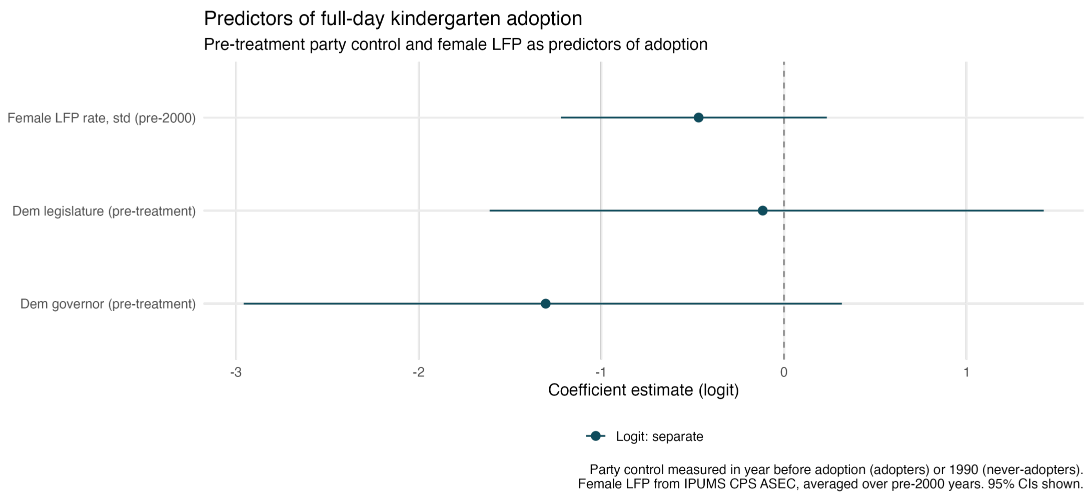

Figure 6.3: Logit predictors of full-day kindergarten adoption

The results are instructive. The coefficients on the Democratic legislature and Democratic governor indicators are both small and statistically indistinguishable from zero, with confidence intervals spanning both positive and negative values, although the Democrat governor seems more negative. This is consistent with the parallel pre-trend evidence in Figure 6.2 and supports the view that the partisan composition of a state government is not a strong predictor of whether or when it adopted a mandatory attendance law. In addition, the coefficient on pre-period female LFP is negative, although imprecisely estimated, which is consistent with the idea that economic demand-side factors are possible predictors of adoption, or that this could arise as a result of a desire to make women work more. This finding aligns with the broader literature, which has similarly found that economic rather than political variables tend to dominate as predictors of FDK adoption timing (Cohen-Vogel et al., 2020; Yang, 2023).

It is important however to not over-interpret these results. The absence of a statistically significant political coefficient does not rule out political influence on adoption, instead it simply suggests that in this cross section the crude indicators of pre-period partisan control do not predict adoption timing. More subtle forms of political endogeneity, such as strategic timing relative to electoral cycles or anticipatory responses to changes in federal funding, cannot be fully ruled out with this analysis. This thus motivates the approach leveraging within-state, district-level variation which is presented in Chapter 7. Indeed, this likely provides stronger identification. By conditioning on district and year fixed effects, it removes some of the confounders in terms of the types of state-level political dynamics that this section documents.

6.3 State-level ITT: political outcomes around adoption

Having documented the evolution of kindergarten enrollment and assessed the degree of political endogeneity in adoption timing, this section studies whether higher kindergarten enrollment intensity is associated with more favorable partisan outcomes for Democrats at the state level. The specification is the TWFE model described in Section 5.1.2, with the three-year rolling average of the kindergarten enrollment share as the main exposure variable, and state and year fixed effects throughout. Figure 6.4 presents the coefficient plot for all six outcome variables considered: partisan lean in the state Senate and state House (the Democratic seat advantage as a share of total seats), the net Democratic seat count in each chamber, a binary indicator for a Democratic governor, and a composite trifecta lean measure. Coefficients are shown for three exposure specifications: the contemporaneous enrollment share, the three year rolling average, and the five year rolling average.

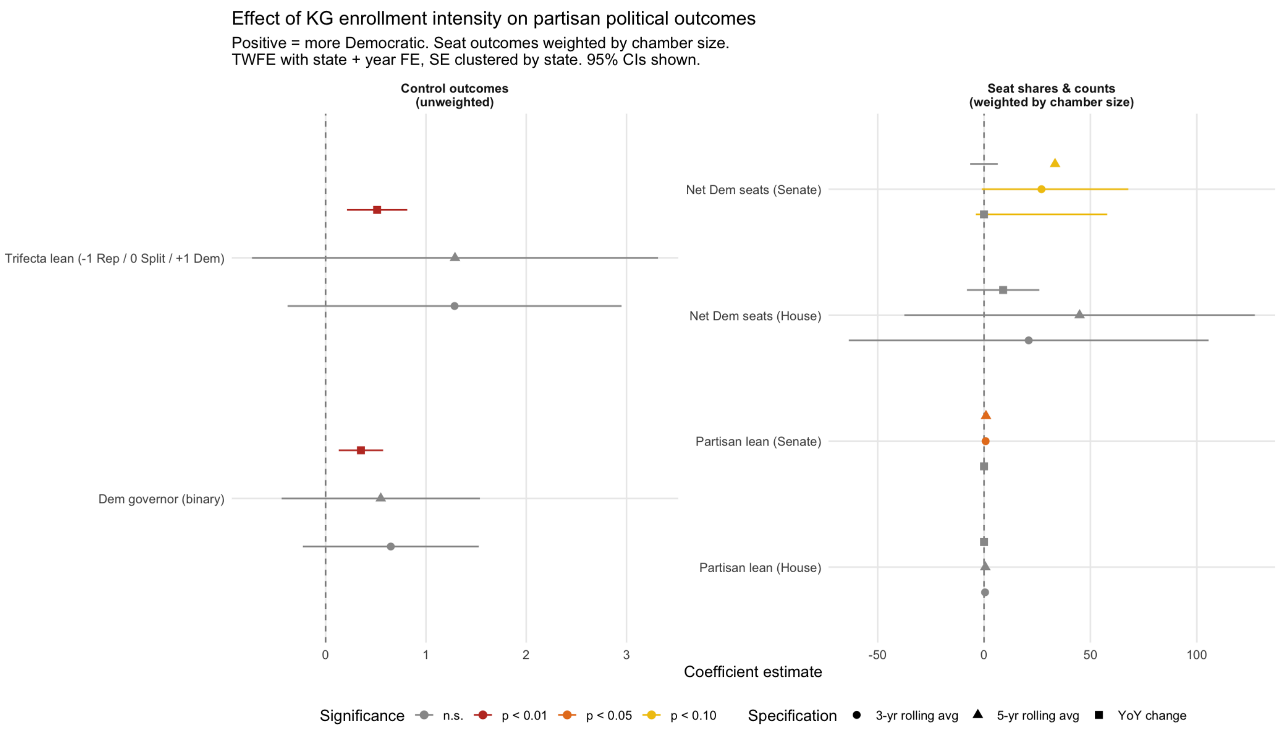

Figure 6.4: TWFE coefficient plot: effect of KG enrollment intensity on partisan political outcomes.

The results suggest a modest positive association between kindergarten enrollment intensity and

Democratic legislative seat shares. The coefficients on the rolling-average exposure for the Senate seats and Senate lean variable are positive, and reach conventional significance thresholds in several specifications, as reported in Table 6.1. In substantive terms, a ten percentage point increase in the share of kindergarten-age children enrolled in kindergarten is associated with a shift of roughly 2.3 to 2.8 extra Democratic seats in the Senate, holding state and year fixed effects constant. A 10 percentage point expansion further increases the likelihood of the Senate leaning Democratic by 8 to 9 percentage points. The year-on-year change specification yields null statistically insignificant estimates for seat counts, which is consistent with the idea that the political effects of kindergarten exposure accumulate gradually rather than manifesting in discrete annual jumps. It may also suggest that the reforms are not necessarily taken in electoral years. House seat estimates are larger in point terms (roughly 3.2 to 4.3 net Democratic seats per ten percentage point increase in enrollment) but are statistically insignificant. These are not very large magnitudes by political standards, but the Senate estimates are consistently in terms of sign, robust across the 3-year and 5-year rolling windows, and statistically distinguishable from zero, which provides suggestive evidence of a compositional effect of early childhood education on state-level partisan outcomes.

The association is considerably weaker and less precisely estimated for executive outcomes. The coefficient on the Democratic governor indicator is positive across specifications but only reaches significance in the year-on-year change specification, as is the trifecta lean measure. This immediate change may either be due to timing, with policies being announced right before gubernatorial elections, or an imprecision of the data. This pattern is consistent with the broader political economy literature, which suggests that highly visible, locally delivered public goods may generate registration and participation dividends (Mani and Mukand, 2007). It may also reflect the fact that state legislative outcomes are more directly connected to school finance decisions, and thus more salient to the families most affected by kindergarten expansion.

Table 6.1: TWFE estimates: KG enrollment intensity – partisan political outcomes

Outcome	3-yr rolling avg	5-yr rolling avg	YoY change
Seat shares & counts (weighted by chamber size)			
Net Dem seats (House)	0.316 (0.373)	0.426 (0.423)	0.034 (0.074)
Net Dem seats (Senate)	0.232* (0.124)	0.283** (0.140)	-0.000 (0.029)
Partisan lean (House)	0.005 (0.003)	0.006 (0.004)	0.000 (0.001)
Partisan lean (Senate)	0.008** (0.004)	0.009** (0.004)	0.000 (0.001)
Control outcomes (unweighted)			
Dem governor (binary)	0.006 (0.004)	0.005 (0.005)	0.004*** (0.001)
Trifecta lean (-1 Rep / 0 Split / +1 Dem)	0.013 (0.008)	0.013 (0.010)	0.005*** (0.002)

Note:
*** p<0.01, ** p<0.05, * p<0.10. Standard errors in parentheses. Partisan lean outcomes weighted by total chamber seats (inverse-variance). Net seat count outcomes are unweighted, each state receives equal influence. State and year FE included. SE clustered by state. Positive coefficients indicate a more Democratic outcome.

6.4 Implications for the causal design

The findings of this chapter provide a mixed but ultimately encouraging picture for the identification strategy developed in Chapter 7. On the one hand, FDK enrollment intensity is associated with modest shifts in Democratic legislative seat shares, which is consistent with the hypothesis that FDK expansion has politically meaningful downstream effects. On the other hand, the endogeneity assessment in Section 6.2 suggests that this association at the state level is unlikely to be driven by the reverse causality of overt partisan political control on adoption decisions, since pre-period partisan composition does not meaningfully predict when states adopted mandatory attendance laws.

However, we cannot extract a causal interpretation of the associations documented in this chapter. Indeed as detailed in Chapter 5: economic, demographic, and fiscal characteristics of states are simultaneously likely to drive both FDK adoption and political outcomes, and the TWFE specification with only state and year fixed effects cannot adequately control for all of these confounders. The results in this chapter should therefore be treated as motivating evidence that the relationship between FDK and political outcomes is worth examining more rigorously, rather than as estimates of the causal effect of interest.

This motivates the shift to the district-level staggered DiD design in Chapter 7. By exploiting variation in the timing of FDK adoption across Pennsylvania school districts, which importantly delegate the decision to local school boards rather than to state legislatures, and by comparing districts that adopted FDK in different years using the Callaway-Sant'Anna framework, the analysis in the next chapter can isolate the effect of FDK expansion from the state-level political and economic dynamics that confound the results here.

7. Analysis and findings II: District/LEA-level causal effects

This chapter presents the main causal results of the thesis. As described in Chapter 5, the identification strategy relies on the staggered adoption of full-day kindergarten across Pennsylvania school districts, combined with the Callaway-Sant'Anna doubly-robust estimator in order to circumvent the biases that would arise when using standard TWFE under treatment effect heterogeneity. This chapter is organized into three sections. Section 7.1 presents the main causal estimates for voter registration, the primary political participation outcome. Section 7.2 turns to electoral outcomes, including vote margins, seat flips, and turnout, which are the most directly policy-relevant question of whether incumbents are electorally rewarded for FDK expansion. Section 7.3 examines heterogeneity across gender and partisan incumbent context, which is central to the mechanism interpretation developed further in Chapter 8.

Before turning to the results, it is worth briefly reviewing the sample on which the estimates are based. The estimation sample consists of 142 treated districts and 53 never-treated districts, with the always-treated group (277 districts) excluded since their pre-treatment period is unobserved. The balance table (Table 4.1) documents pre-treatment means by group. A few patterns emerge. Districts that get treated early (those adopting before the median treatment year) have a notably lower pre-treatment registration rate (0.0092) than both those that get treated late (0.0158) and never-treated districts (0.0197). They are also smaller on average, with a mean population aged 25–44 of approximately 6,559, compared to 7,628 for those treated late and 8,671 for never-treated districts. These differences in levels are precisely what district fixed effects are designed to absorb, and they reinforce the importance of using a within-district estimator rather than a cross-sectional comparison. The female share of the adult population is very similar across all three groups (roughly 0.511–0.513), which is reassuring as a balance check on the demographic dimension most relevant to the mechanism analysis. Importantly, this section focuses on adults aged 25-44, as it seems to be the most relevant group to study when studying the populations affected by kindergarten expansion. This assumes the moderately conservative bounds of 20 and 39 years old as the age at which most people have their children, as we are studying those with 5 year olds. However, the 277 always-treated districts, more numerous than the 142 treated districts themselves, are entirely excluded from the DiD analysis by design, since their pre-treatment period is unobserved. The causal estimates therefore apply specifically to districts that transitioned to FDK during the sample window, which tend to be of moderate fiscal capacity and population size. The never-treated districts, by contrast, are on average considerably wealthier and larger than treated districts, as documented in Table 4.1. The estimates should not be extrapolated to either tail of the district distribution without caution. The descriptive discrepancies motivate the use of the not-yet-treated districts as the control group, in order to maximize comparability with the treated districts.

7.1 Main causal estimates: voter registration

7.1.1 Primary event study

Figure 7.1 presents the primary result of the causal analysis: the Callaway-Sant'Anna event-study plot for the new voter registration rate among adults aged 25–44, using not-yet-treated districts as the control group. Each point on the plot represents the average treatment effect on the treated (ATT) for a given year relative to FDK adoption, where $t = 0$ denotes the year of the switch and negative values correspond to pre-treatment periods.

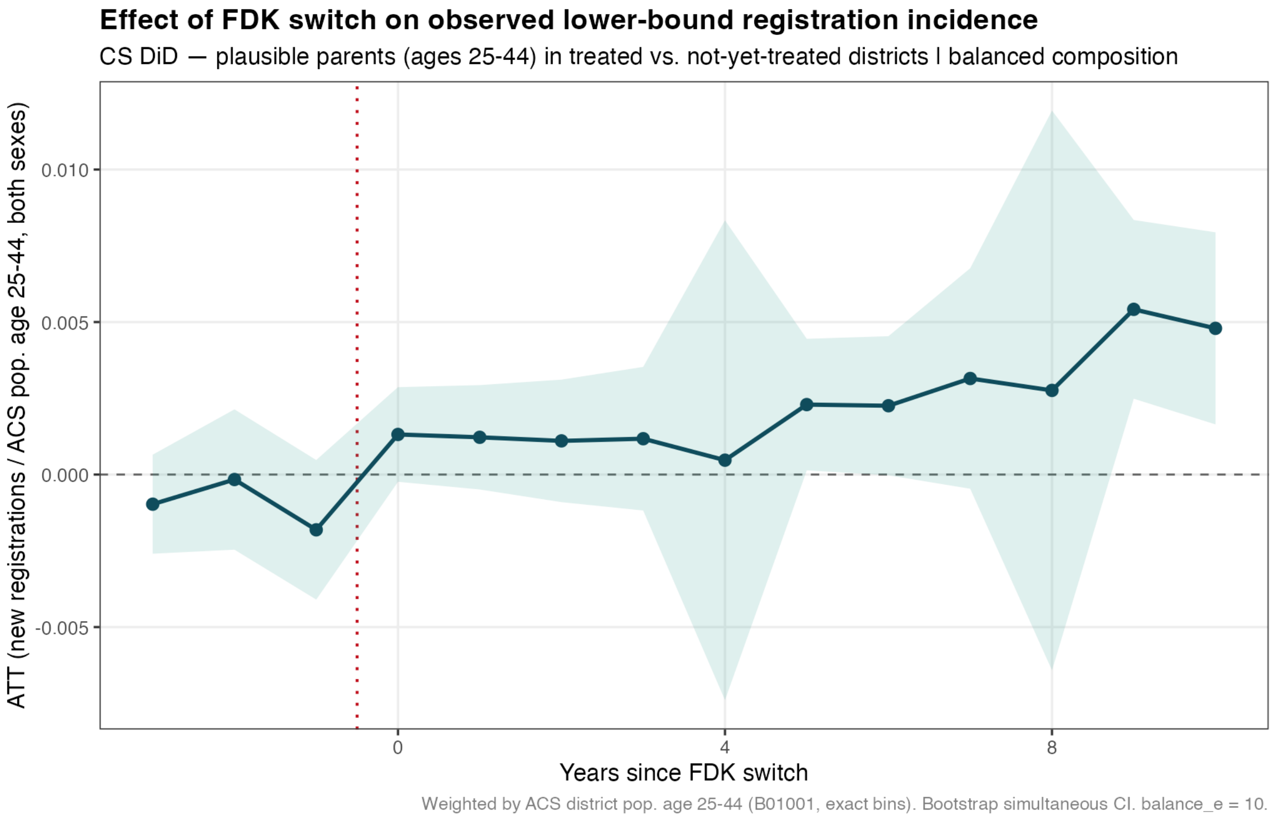

Figure 7.1: Primary CS DiD event study: ATT on new voter registration rate, adults 25–44

The pre-treatment period highlights dynamics which are consistent with the necessary identifying assumptions. The estimates for $t < 0$ are close to zero and do not display any systematic pre-trend, either upward or downward. This is consistent with the parallel trends assumption underlying the CS DiD design, and suggests that treated and not-yet-treated districts were on similar registration trajectories in the years leading up to FDK adoption, despite the possibly different levels identified in the balance table.

Following adoption, the event study shows a positive and persistent shift in the registration rate. The effect is somewhat immediate at $t = 0$, albeit not statistically significant. It remains positive but insignificant, and rises slowly over time after $t + 4$, reaching an increase of roughly 0.5 percentage points in the registration rate, statistically different from zero at $t + 9$ and $t + 10$. A sharp instantaneous spike at $t = 0$ or $t = +1$ would be consistent with a pure salience or credit-claiming channel, whereby parents register in direct response to a salient policy announcement, and although we do see a spike it is not necessarily what could be considered sharp. The more gradual accumulation observed here is instead more consistent with a mechanism operating through changes in time constraints or labor market behavior, which could be expected to materialize progressively as families adjust to the new childcare arrangement. This point is returned to in more detail in Chapter 8 in the context of the labor force participation results, notably the maternal ones. Here, as well as in other future uses of this specification, for $t < 1$, 84 not-yet-treated districts are used as controls, for $t \in \{2; 3; 4; 5\}$ 78, then for each year after it decreases to 67, 52, 44, 44, and finally 39 at $t = 10$. Figure A.4 in the Appendix visually shows the evolution of the number of control groups.

7.1.2 Scalar ATT and specification robustness grid

The simple ATT, which averages the post-adoption estimates across all treated cohorts and years weighted by cohort size, is reported in Table 7.1. For the pooled specification (adults aged 25 to 44, doubly-robust estimator, not-yet-treated districts as the control group) the estimated ATT is 0.00330, and it is significant at the 1% level. To put this in context, the pre-treatment mean registration rate in treated districts is approximately 0.009 to 0.016 depending on cohort (see the balance table), so an ATT of 0.00330, corresponding to a yearly increase of 0.33 percentage points, represents a meaningful relative increase of roughly 20 to 35 percent relative to the pre-treatment baseline. This is not a trivial magnitude, though as with all estimates of this type it is important to keep in mind that the registration rate is a lower bound on the true new-registration incidence, so the absolute level should be interpreted with caution. This table moreover shows a stronger effect for fathers than for mothers. While this result may seem counterintuitive, the related mechanism will be further explored in the heterogeneity section. Moreover, considering the political outcomes the Democrat ATT (0.00080, SE 0.00029) is significant at the 1% level, whereas the Republican ATT (0.00154, SE 0.00084) has a larger point estimate but is only significant at the 10% level. The standard errors differ because the Republican registration rate has a higher baseline variance, thus the correct test of partisan asymmetry is the net Dem - Rep ATT (-0.00074). As it is insignificant, this reaffirms a lack of systematic advantage for one party over the other.

Table 7.1: CS DiD estimates: FDK adoption on voter registration

Outcome	ATT (SE)
Registration rate (pooled)	0.00330*** (0.00124)
Registration rate (mothers)	0.00300** (0.00125)
Registration rate (fathers)	0.00359*** (0.00123)
Democrat registrations	0.00080*** (0.00028)
Republican registrations	0.00154* (0.00086)
Net Dem minus Rep registrations	-0.00074 (0.00081)

Note:
*** p<0.01, ** p<0.05, * p<0.10. Standard errors in parentheses. CS DiD DR estimator, not-yet-treated control. Weighted by ACS pop. 25-44. Positive = more Democratic.

A compositional caveat applies to these late-horizon estimates. As visible in Figure A.5, the 142 treated districts are spread across adoption cohorts from 2007 to 2023, with the largest waves concentrated in 2009 (20 districts) and 2010 (18 districts). In a panel balanced at e = +10, only cohorts treated in 2015 or earlier contribute to the t + 10 estimate, and only those treated in 2016 or earlier contribute to t + 9.

This corresponds, respectively, to 86 and 91 districts. The later and more numerous cohorts, including the 15 districts switching in 2018 and the 11 in 2019, are observed only in the early post-adoption window and drop out of the event study at longer horizons. The statistical significance at t + 9 and t + 10 should therefore be interpreted with some caution: it reflects the experience of the earliest-adopting cohorts specifically, which may differ from later adopters in ways that also predict registration trends. The scalar ATT reported in Table 7.1, which aggregates across all cohorts and post-adoption years, is less sensitive to this composition issue and is the preferred summary of the overall effect.

Figure 7.2 presents the robustness grid, showing the event-study path across three alternative specifications: unweighted doubly-robust with never-treated control, doubly-robust with not-yet-treated control, and regression-adjusted with not-yet-treated control. The four paths track each other closely in both the pre-period and the post-period (save for the unweighted one in the pre period, but this is not surprising). The not-yet-treated specification produces a slightly smaller point estimate in the post-period but remains positive and statistically distinguishable from zero. This consistency is important because it addresses the concern that not-yet-treated districts are structurally unusual in a way that would make them a poor counterfactual for the treated districts. The fact that the never-treated specification produces similar results suggests that the positive ATT is not simply due to the choice of control group.

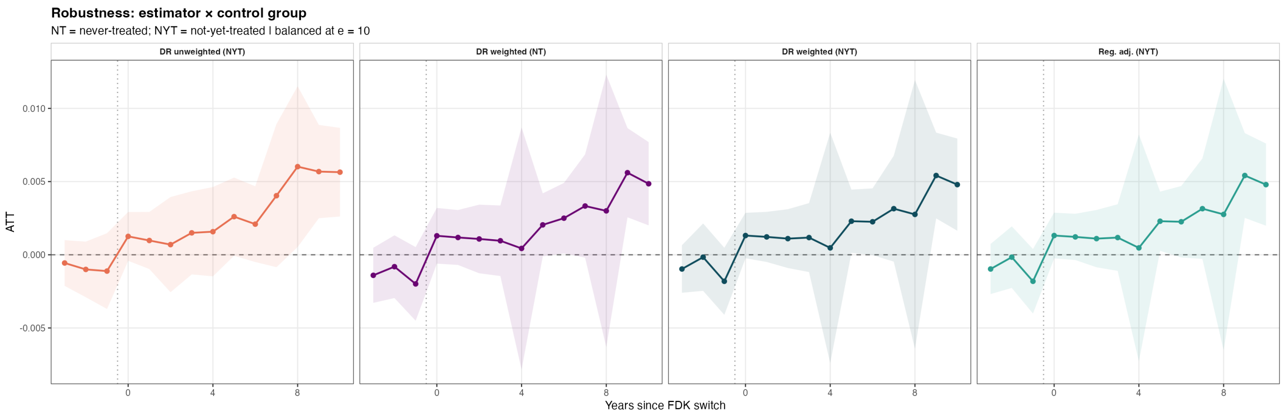

Figure 7.2: Robustness grid: event-study paths across three specifications. NT = never-treated; NYT = not-yet-treated. Shaded bands = bootstrap simultaneous 95% CIs. Consistency across panels supports the main finding.

As an additional check, Figure 7.3 overlays the standard TWFE event study with the CS DiD estimates for the pooled sample, which makes salient the magnitude of the heterogeneity bias that would result from the conventional estimator. The two series diverge in the later half of the post-period. the TWFE estimates are attenuated relative to CS DiD, which is the classic symptom of negative-weighting bias under staggered adoption. This comparison is not merely illustrative, indeed it demonstrates that in this specific application, the use of standard TWFE would lead to substantively misleading conclusions, and motivates the choice of the CS DiD estimator as the benchmark throughout. This further validates the robustness checks performed in Chapter 9.

Overall, the evidence at the pooled level suggests that expanding full-day kindergarten has a positive impact on registration to vote.

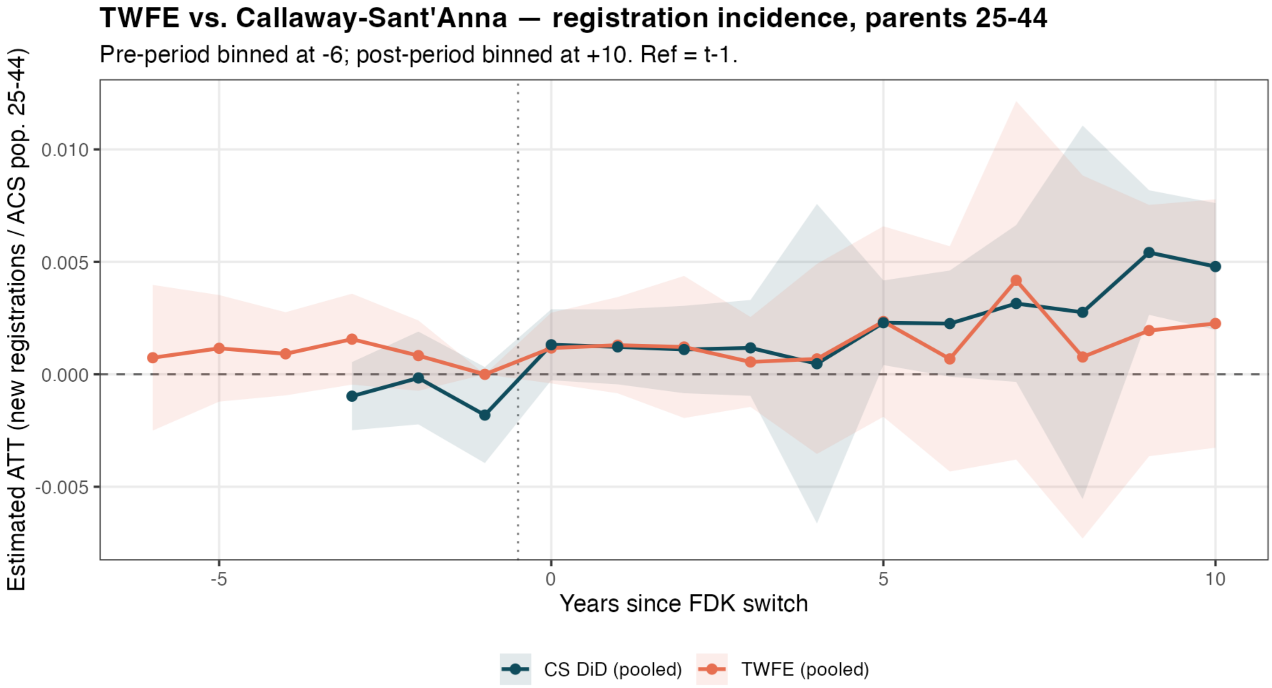

Figure 7.3: TWFE (orange) versus CS DiD (teal) event-study comparison

7.2 Main causal estimates: electoral outcomes

The voter registration results establish that FDK adoption increases civic engagement as measured by new voter registration. The more directly policy-relevant question, however, is whether this translates into electoral benefits for the politicians responsible for the expansion. This section therefore turns to outcomes measured at the legislative district level: vote margins, incumbent retention probabilities, turnout, and seat flips. The most micro level that could be responsible for this change is the school board level, but this presents a few disadvantages. Firstly, members are not always associated with a political party, and depending on the size of the districts, interpersonal relationships could shape electoral outcomes more than policies. Second, and most importantly, the data on school board composition is not readily available for all 500 districts, and the member by member composition may be driven by individual level factors (ex: choosing to retire, health issues, ...) which would provide less generalizable results. In this light, this thesis instead focuses on legislative outcomes such as state Senate and state House elections. This is made possible as the voter file also matches every individual with their assigned congressional districts, and Klarner vote data includes information on state level legislative results.

7.2.1 Estimation design for electoral outcomes

Studying electoral outcomes introduces an additional layer of complexity relative to the registration analysis. School districts and legislative districts (Pennsylvania House and Senate constituencies) do not share boundaries, which requires a mapping between the two. Each school district is assigned to its most common legislative district based on the addresses of registered voters in the district. Treatment is then defined at the legislative district level as the first election cycle after FDK adoption in the majority of school districts contained within it. The overlap is however far from perfect, due to different numbers of each type of district. For clearer insight on the geographical boundaries, see Section A.1.

For House districts, the same CS DiD framework is applied, but time is now measured in election cycles rather than calendar years: $t = 0$ is the first even-year election following FDK adoption, and $t = -1$ is the preceding election two years earlier. Because Pennsylvania House elections occur every two years, there are typically three to five pre-treatment election cycles available for parallel trends assessment, and four post-treatment cycles within the observation window. The not-yet-treated group serves as the control here rather than the never-treated, since the latter are too few and potentially structurally different at the legislative district level.

Senate districts present an additional identification constraint. Pennsylvania only has 50 state Senate districts, compared to 203 House districts, and their larger geographic footprint means that very few are cleanly matched to a single treatment cohort, making the CS DiD approach impractical. The Senate results therefore use the intensity-of-treatment two-stage imputation estimator of Borusyak et al. (2024), applied in the spirit of Gardner (2022) to the continuous intensity setting. In the first stage, unit and time fixed effects are estimated on untreated observations only (that is Senate districts observed in election cycles prior to their first FDK adoption) to recover an imputed counterfactual $\hat{Y}_{dt}(0)$ for all district-year cells. In the second stage, the residual $Y_{dt} - \hat{Y}_{dt}(0)$ is regressed on the interaction $\text{Intensity}_d \times \text{Post}_{dt}$, where $\text{Intensity}_d \in [0, 1]$ is the time-invariant share of the Senate district's registered voters residing in treated school districts, and Post_{dt} is an indicator for election cycles at or after the district's first treatment year. The coefficient on this interaction is interpreted as the effect of moving from zero to full FDK exposure within a Senate district. An event-study version is also estimated, replacing the single post indicator with a set of relative-election-cycle dummies interacted with Intensity_d, with $t = -1$ omitted as the reference period. Standard errors are clustered at the district level throughout.

7.2.2 Event-study evidence

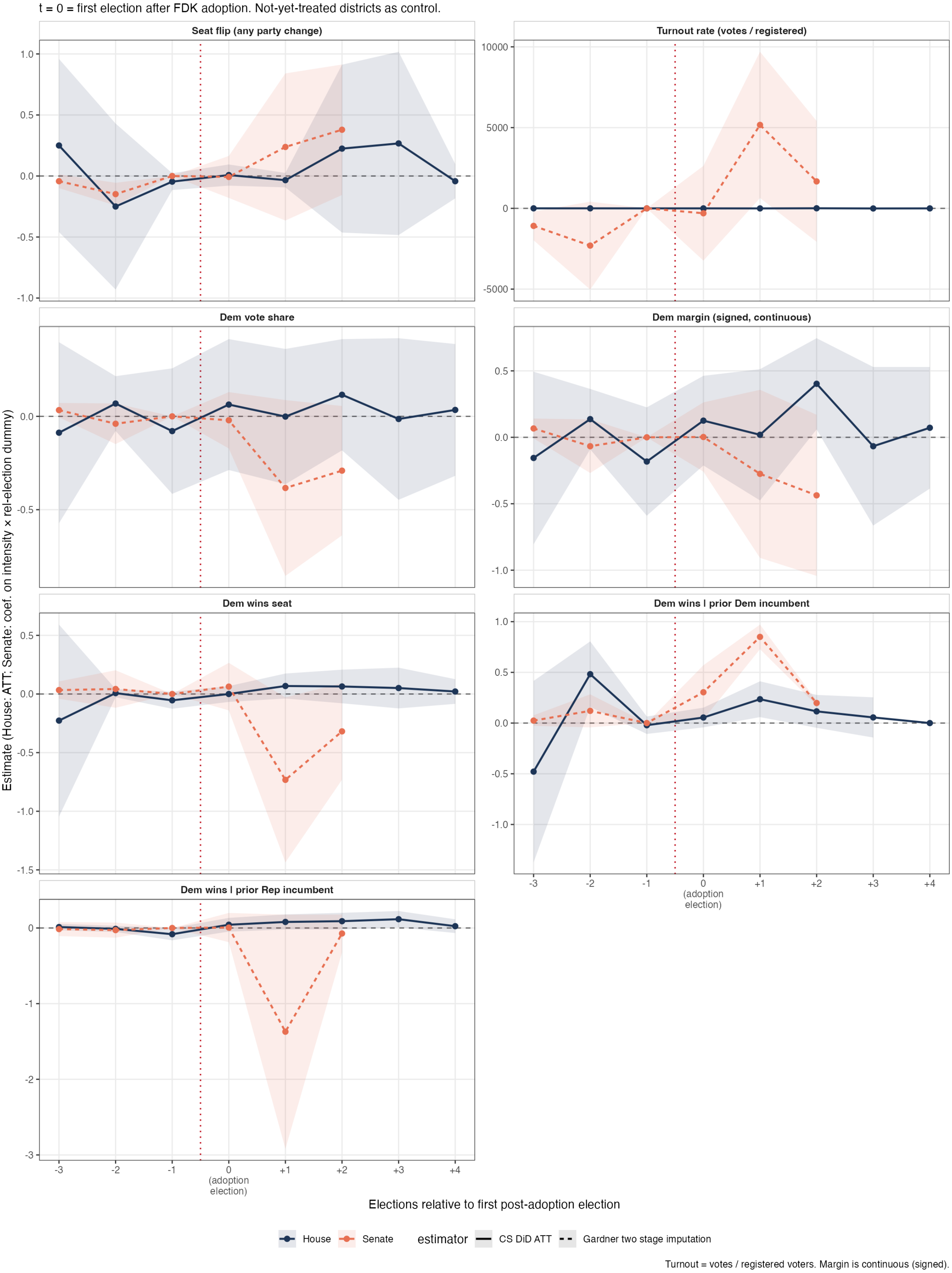

Figure 7.4: Electoral outcome event studies

Figure 7.4 presents the full event-study grid for all seven electoral outcomes: Democratic vote share, Democratic win probability, probability of a Democratic win conditional on a prior Democratic incumbent, probability of a Democratic win conditional on a prior Republican incumbent, the signed Democratic vote margin, turnout rate, and probability of a seat-party flip. House results are shown as solid lines (CS DiD) and Senate results as dashed lines (Gardner two stage imputation). All series are indexed to $t = 0$, the first election cycle after FDK adoption.

Several patterns emerge from the full grid. First, the pre-election-cycle trends are broadly flat across outcomes, which provides support for the parallel trends assumption in election time, and is a reassuring check given that the electoral panel is considerably shorter and noisier than the registration panel. Second, the estimates are generally imprecisely estimated across outcomes and chambers, with wide confidence intervals that encompass zero for most post-adoption periods. This is not entirely surprising: with 203 House districts and an even smaller number of usable Senate districts, the effective sample size for the electoral analysis is substantially smaller than for the registration analysis, and the binary nature of several outcomes (win/loss) reduces statistical power further. Despite this imprecision, a few directional patterns are worth discussing. Figure 7.5 presents the focused subset of outcomes discussed in the subsection below.

7.2.3 Incumbency and vote margin

Figure 7.5 focuses on the three outcomes most directly relevant to the electoral interest of FDK-expanding incumbents: the probability of a Democratic win conditional on a prior Democratic incumbent, the probability of a Democratic win conditional on a prior Republican incumbent (that is, a district classified as Republican by the proxy), and the signed Democratic vote margin.

The conditional win probabilities can highlight mechanisms through which FDK expansion could translate into electoral benefits. If FDK expansion generates a credit-claiming dividend for the incumbent party, one would expect Democratic incumbent retention rates to rise following adoption. If instead FDK mobilizes new Democratic or Republican registrants who then vote at high rates, one might observe challengers defeating opposite party incumbents at higher rates. If it is a more general civic mobilization effect, neither of these partisan patterns should dominate.

Graphical patterns for some House outcomes are not fully aligned with the scalar ATT estimates. Given this instability, the House electoral evidence should be treated as mixed rather than directionally favorable to either party. Indeed, the ATT for these categories is negative (see Table 7.2), and there is a statistically significant decrease in the vote margin. The effect is positive but insignificant for a Democrat win with a Democrat incumbent, and negative but insignificant in the case of a Republican incumbent, which perhaps suggests a rewarding effect at the House level. Importantly, the House districts are smaller so may more closely match the school district boundaries (although of course imperfectly, since there are 203 state representatives, and 500 districts). The state Senate results are roughly similar, and may be more consistent with a credit claiming mechanism. The Democratic margin statistically significantly shrinks following a reform, and Democrats have an insignificant but pointwise negative win probability in the case of a Republican incumbent. In the case of a Democrat incumbent however, the effect is positive and statistically significant at the 1% level, and associated

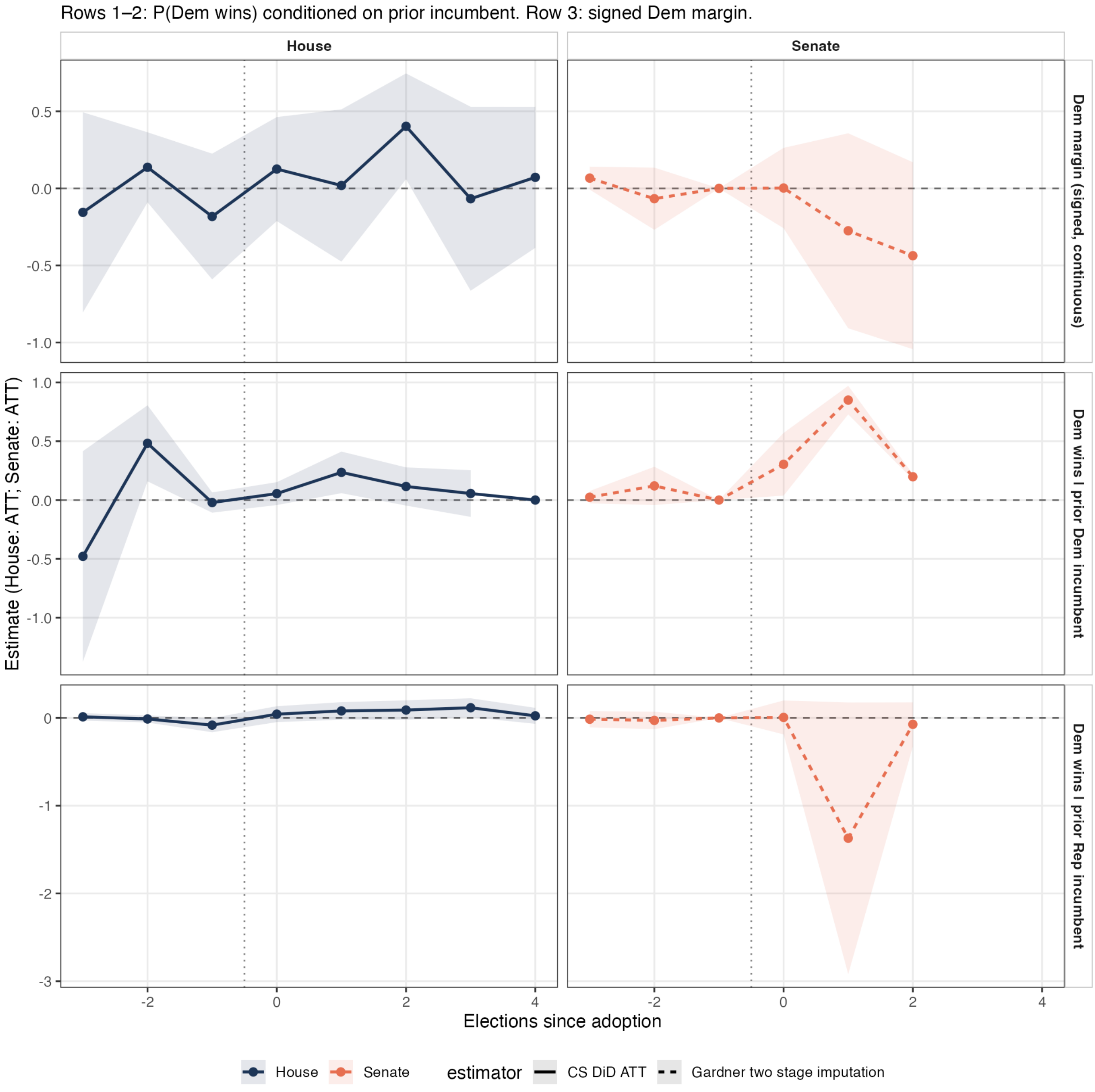

Figure 7.5: Focused electoral outcomes

with a large ATT of 0.249. However, treatment is imperfectly defined at this level due to the small sample size and the geographical mismatch in boundaries. It is thus complicated to properly extract conclusions from these figures. The most appropriate outcome to look at would be school district elections, however this is difficult to implement and plagued by many other biases as described above.

Table 7.2: Electoral outcomes

Outcome	House ATT (SE)	N treated	N control	Senate ATT (SE)	N districts
Seat flip (any party change)	-0.0381 (0.0435)	87	47	0.2107 (0.1354)	50
Turnout rate (votes / registered)	-2.8562*** (0.9295)	87	47	2594.8953* (1437.4335)	50
Dem vote share	-0.0573* (0.0329)	87	47	-0.2277** (0.0967)	50
Dem margin (signed, continuous)	-0.1018* (0.0554)	87	47	-0.2476* (0.1417)	50
Dem wins seat	-0.0602 (0.0432)	87	47	-0.3090** (0.1434)	50
Dem wins | prior Dem incumbent	0.0495 (0.0823)	68	36	0.2496*** (0.0668)	37
Dem wins | prior Rep incumbent	-0.0127 (0.0385)	72	47	-0.1714 (0.1317)	35

Note:

*** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at district level.

House: CS DiD DR estimator, not-yet-treated control, election-cycle time.

Senate: Gardner DiD estimator, not-yet-treated control, election-cycle time. Dem margin = (Dem - Rep) / total.

The electoral results warrant more direct engagement than a simple appeal to statistical power. The arithmetic linking the registration result to vote shares is tractable and informative. The estimated ATT of 0.33 percentage points in new registrations per year, applied to a mean adult population aged 25–44 of approximately 7,000 in treated districts, is consistent with roughly 23 additional new registrations per district per year. Compounded over the ten-year post-adoption window over which the effect accumulates, this corresponds to approximately 230 additional cumulative registrants per district. Pennsylvania House districts contain on average around 60,000 to 70,000 registered voters, and competitive races in the state have frequently been decided by margins of 500 to 2,000 votes. Under the optimistic assumption that all 230 new registrants turn out at the same rate as existing voters and split their votes proportionally to the existing district composition, the implied vote-share shift is on the order of 0.3 to 0.4 percentage points per district, a small but not negligible point estimate in a close race.

The electoral results in Table 7.2 tell a more complicated story than a simple null. At the House level, the Democratic vote share ATT is -0.057^{*} and the Democratic margin ATT is -0.102^{*}, both negative and significant at the 10% level. At the Senate level, the Democratic vote share (-0.228^{**}), Democratic margin (-0.248^{*}), and Democratic win probability (-0.309^{**}) are all negative and significant, suggesting that FDK adoption in a district is if anything associated with worse Democratic performance in the aggregate. The single exception is the Senate result for Democratic wins conditional on a prior Democratic incumbent, which is positive and significant at the 1% level with an ATT of 0.250. Taken together, this pattern may be consistent to some extent with a general credit-claiming dividend for the party associated with FDK expansion, given that the majority of switching districts are classified as Republican-leaning under the pre-reform registration-majority proxy (95 out of the 142).

Three explanations for this pattern should be considered. The first is boundary mismatch attenuation. School districts that adopted FDK during the sample window tend to be of moderate fiscal capacity and are not randomly distributed across legislative constituencies. If FDK-adopting school districts are disproportionately located in legislative districts that are trending Republican for reasons unrelated to the policy, the negative aggregate coefficients could reflect compositional sorting rather than a causal effect of FDK on Democratic performance. The second is non-partisan mobilization. As will be shown in Section 7.3.2, the registration gains documented in Section 7.1 are non-partisan, with no statistically significant net Democratic advantage in new registrations. If newly mobilized voters split proportionally or slightly toward Republicans in these districts, a negative aggregate Democratic margin would be consistent with the registration result rather than contradicting it. The third is that the positive Senate result for Democratic incumbent retention reflects a genuine but narrow credit-claiming effect: Democratic incumbents who can credibly claim association with FDK expansion benefit electorally, while the broader electorate shifts in the other direction. This interpretation is speculative and the Senate result should be treated cautiously given the small sample of 50 districts, the imperfect boundary mapping, and the fact that several other Senate outcomes point in the opposite direction.

The registration results, which are based on a larger and more continuous panel with a cleaner treatment definition, should be considered the primary evidence throughout. The electoral results are suggestive but too noisy and too dependent on imperfect geographic mappings to support strong causal claims in either direction.

7.3 Heterogeneity: voter registration

7.3.1 By gender

The most theoretically motivated heterogeneity analysis concerns the distinction between mothers and fathers. As discussed in Chapter 2, the time-reallocation mechanism predicts that FDK should have a stronger effect on mothers, who face more binding childcare constraints on average and who are therefore most likely to gain slack in their daily time allocation when a kindergarten expansion provides all-day childcare. A strong concentration of the registration effect among mothers relative to fathers would constitute evidence for this channel; roughly equal effects would instead point toward a more diffuse household-level socialization mechanism. Figure 7.6 plots the event-study estimates separately for potential mothers (females aged 25–44, using the female ACS denominator from B01001) and potential fathers (males aged 25–44). The corresponding simple ATTs are 0.00300 (SE 0.00124, $p<0.05$) for mothers and 0.00359 (SE 0.00126, $p<0.01$) for fathers.

The results do not support the time-constraint prediction. The father ATT is larger than the mother ATT in point terms, and while the difference is within sampling error, its direction is the opposite of what a maternal childcare-constraint channel would predict. The registration effect is gender-neutral, and this finding directly contradicts the narrow version of the labor force participation mechanism. The most coherent interpretation of this pattern is a household-level socialization channel: FDK relaxes a constraint at the household level rather than specifically a maternal one, and both parents benefit from the resulting reallocation of time and community exposure in similar measure. Crucially, the timing of the effect reinforces this reading. A time-reallocation mechanism, maternal or otherwise,

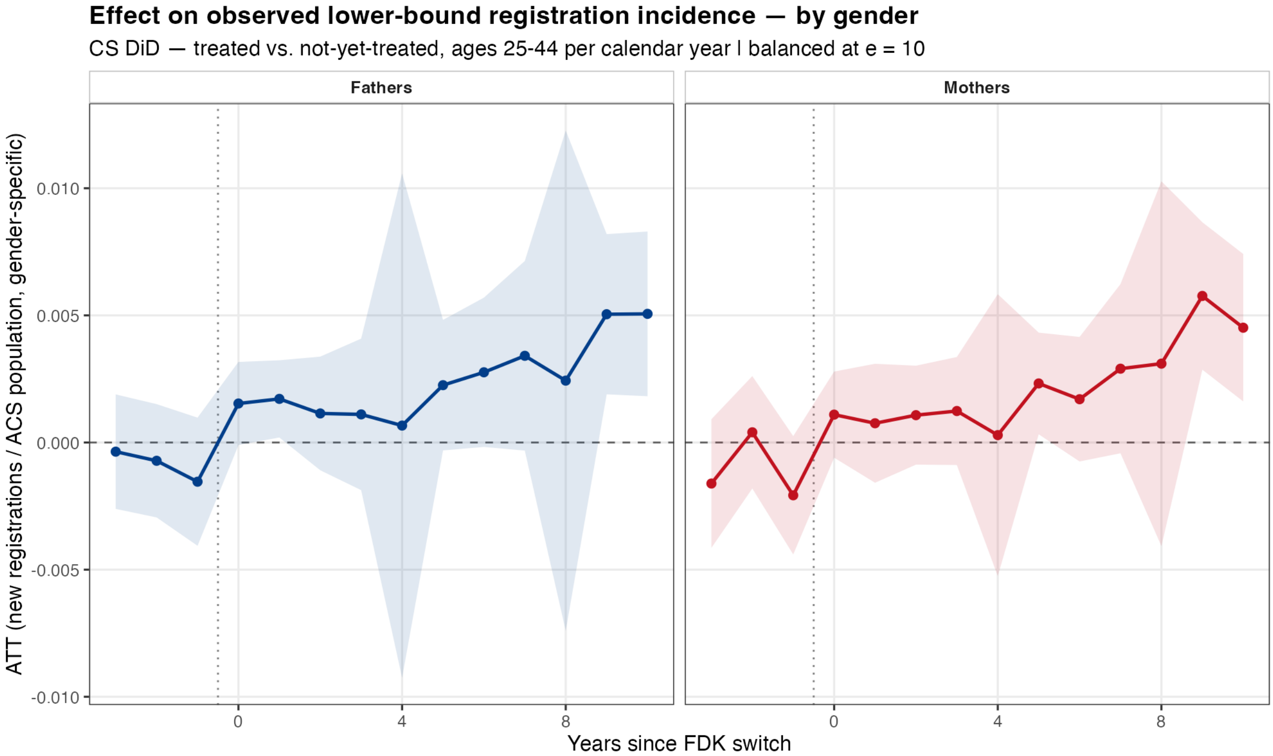

Figure 7.6: Registration effect by gender

would predict a registration response concentrated at t=0 to t=+1, as parents adjust immediately to the new childcare arrangement. The event-study paths for both mothers and fathers instead show a slow and gradual post-adoption build-up that reaches statistical significance only at t=+9 and t=+10 under simultaneous confidence bands. This is inconsistent with any immediate behavioral response, whether driven by time constraints or salience, for either gender.

The labor force participation results in Chapter 8 introduce a further tension rather than resolving it. The LFP effect documented there is concentrated among single mothers of kindergarten-age children, which is consistent with the maternal time-constraint story at the labor market margin. However, if that mechanism were the primary driver of civic engagement, one would expect the registration effect to be correspondingly larger for mothers than for fathers, which it is not. These two findings point in different directions and cannot be fully reconciled with the data available. The most defensible conclusion is that the registration effect reflects a gradual household-level integration process that benefits both parents, while the LFP effect captures a distinct and more targeted labor market response among the most constrained households. Whether and how these two channels interact is a question that individual-level data linking employment status directly to registration records would be needed to answer, and it remains an open direction for future research.

7.3.2 By party registration

Another key dimension of heterogeneity to consider is, irrespective of incumbency, which party gains the most. Figure 7.7 presents the event-study estimates for Democratic and Republican new registrations separately, and Figure 7.8 shows the net Democratic minus Republican registration rate. The corresponding ATTs are 0.00080*** for Democratic registrations and 0.00154* for Republican registrations. The net Democratic minus Republican ATT is −0.00074. (See Table 7.1)

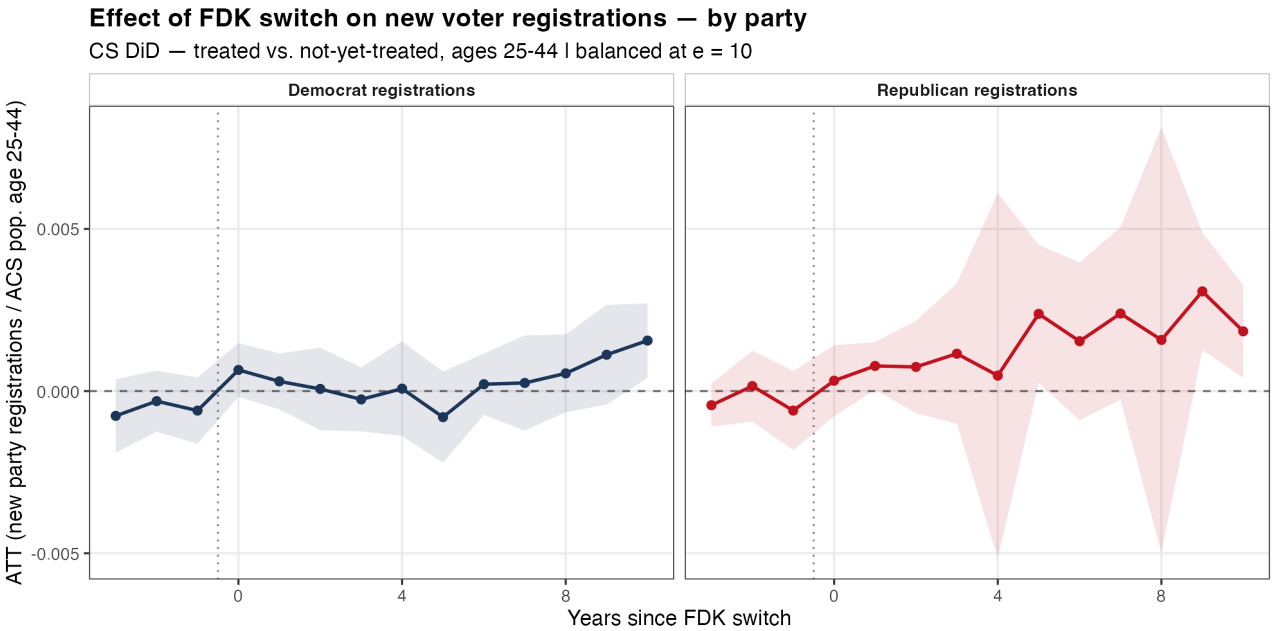

Figure 7.7: Effect on new voter registrations by party

These results require careful interpretation. At first glance, the fact that the Democratic registration ATT is more statistically significant than the Republican registration ATT might suggest a partisan mobilization effect in favor of Democrats. However, the net Dem-minus-Rep ATT is itself negative (though not significantly so), which means that the larger point estimate for Republican registrations, despite its statistical imprecision, offsets any apparent Democratic advantage when the two series are differenced. The confidence intervals on both individual-party outcomes and the net outcome are wide enough that strong conclusions about relative partisan advantages are not warranted. Moreover, the figures make salient that the point estimates for new Republican registrations are systematically higher, but this does not yet look at results differentiated by party incumbency, something that will be further explored in the next section. A more parsimonious interpretation is that FDK adoption generates a broad civic mobilization effect that draws in new registrants from both parties.

This null finding on partisan mobilization is itself informative. It suggests that FDK expansion does not generate a systematically one-sided registration dividend, at least at the aggregate level. This is consistent with the broader pattern in the electoral outcomes section: the effects on vote margins and incumbent win probabilities, while directionally negative for Democrats, are imprecise and should not be interpreted as conclusive evidence of partisan tilting.

7.3.3 By incumbent party

A final heterogeneity dimension concerns the partisan context of the district at the time of FDK adoption. The prediction from the credit-claiming literature (Mani and Mukand, 2007) is that the registration dividend, if any, should be more Democrat leaning in districts with Democratic incumbents, who might be more closely associated with an expansion of public services. In Republican-incumbent districts, we may either see the same effect or alternatively families may not attribute the FDK expansion to the incumbent, which would dampen any electoral reward. Although this paper is unable to know the

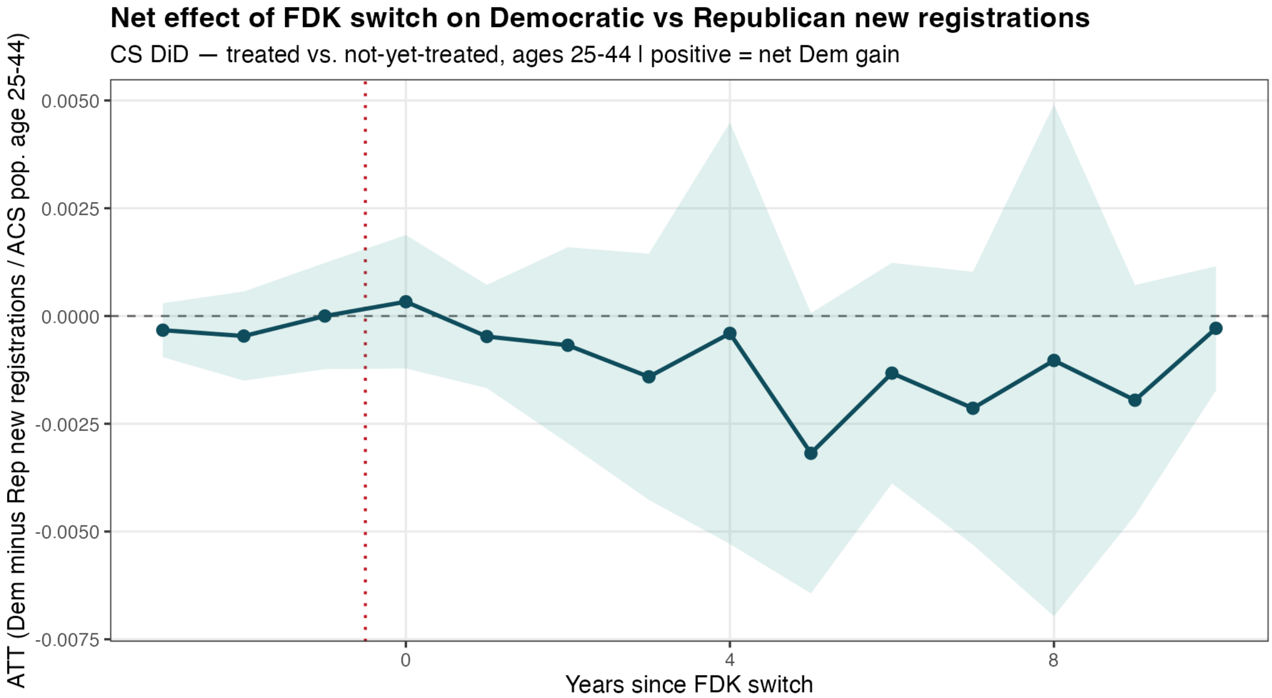

Figure 7.8: Net Democratic minus Republican new registration rate

exact incumbency composition, a proxy is constructed which corresponds to the most represented party among people registered to vote prior to the reform. Figure 7.9 plots the event-study estimates for the registration rate, separated by whether the district had a Democratic or Republican registration majority at the time of FDK adoption.

The results suggest that the registration effect is present in both types of districts, though the confidence intervals are wide given the reduced sample sizes when conditioning on incumbent party. The point estimates are broadly similar in magnitude for each party irrespective of the incumbency status. Importantly though, of the 142 switching districts, 95 are initially Republican, 45 are initially Democrat (and 2 are missing). Thus due to the geographical imprecision in matching to House districts to obtain incumbency status, the lack of significance and noisiness of the results is to be expected. If we take the more conservative interpretation that there are no significant gains for either party even conditioning on incumbency, this is consistent with the view that FDK expansion generates a general civic participation dividend rather than a targeted partisan mobilization. It may also reflect the fact that school boards in Pennsylvania often run without formal party labels, which would weaken the attribution mechanism. In this light, parents could register in response to the FDK policy without clearly attributing it to a Democratic or Republican official.

Overall, the heterogeneity results paint a picture of FDK expansion as a broadly civic-mobilizing policy, with effects on registration that are present across gender, party, and incumbent context, and that do not strongly concentrate in any one subgroup. Chapter 8 examines the labor force participation channel as one candidate mechanism, but as will be seen, the LFP evidence is concentrated among single mothers while the registration effect is gender-neutral, a tension that the available data cannot fully resolve.

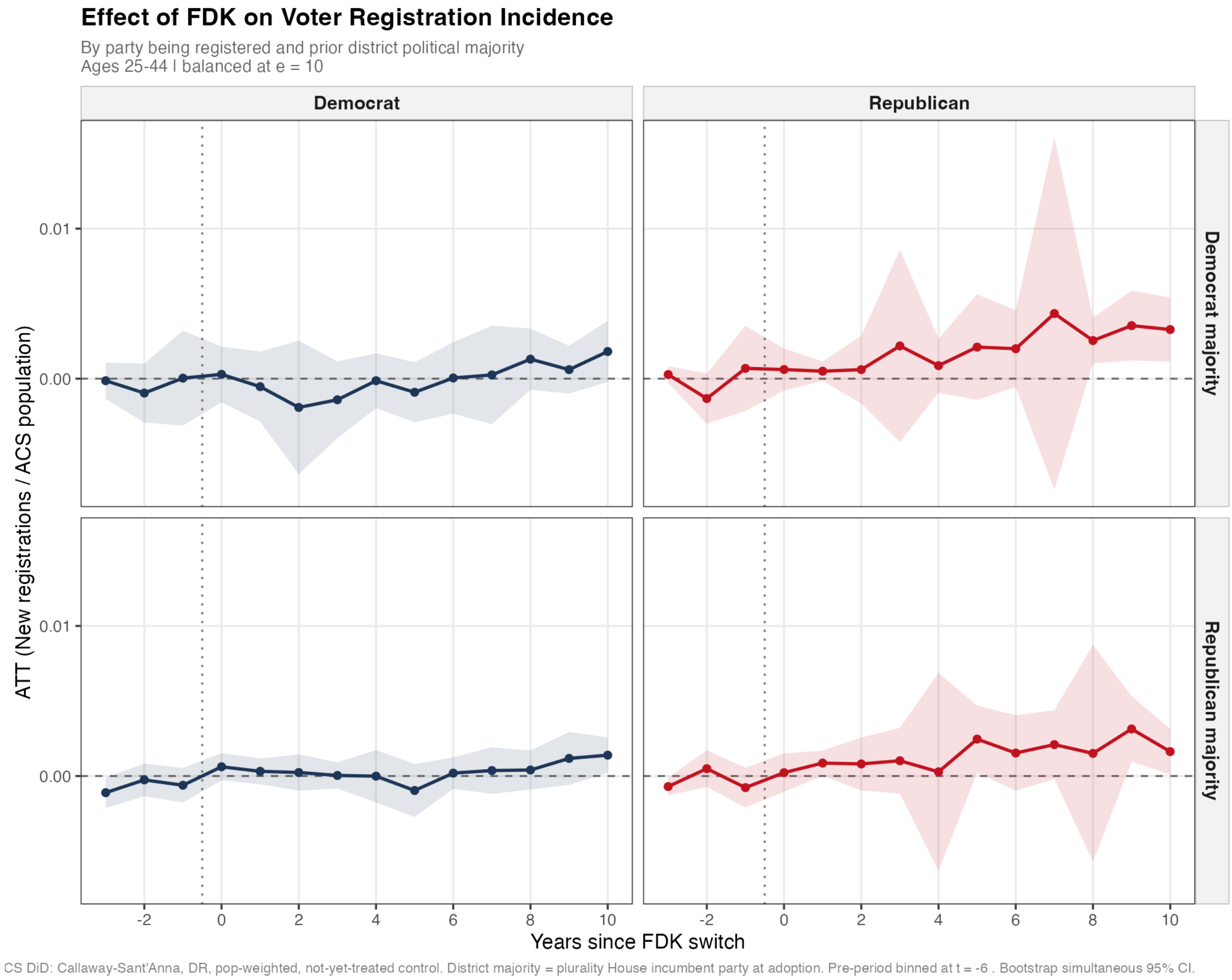

Figure 7.9: Heterogeneity by incumbent party: effect of FDK adoption on voter registration rate

8. Mechanism: Labor force participation

The previous chapter documented a positive and statistically significant effect of full-day kindergarten adoption on new voter registration rates, with an ATT of 0.33 percentage points for the pooled sample, robust across estimator specifications and control group choices. The event-study path showed a gradual post-adoption build-up rather than an immediate spike, which argued against a pure salience or credit-claiming channel and pointed instead toward a mechanism that materializes progressively over time. Critically, Section 7.3.1 established that the registration effect is gender-neutral: the ATT for fathers (0.00359) is marginally larger than for mothers (0.00300), a pattern that directly contradicts the narrow maternal time-constraint channel. This chapter investigates the labor force participation channel not as a full explanation of the registration result, but as a test of whether a labor market response is occurring at the margin where the childcare constraint is most binding, among single mothers, even if that channel cannot account for the broader gender-neutral pattern in registrations.

If FDK functions primarily as a free all-day childcare substitute, it directly relaxes the time constraints of parents, and most acutely those of single mothers, who face that constraint without a second adult to share childcare duties. A relaxation of this constraint could then increase participation in the labor market, which in turn increases exposure to colleagues, communities, and social networks, and may ultimately translate into higher civic engagement and propensity to register to vote. This chapter tests whether the labor market channel is empirically active at this margin, with the prior expectation given the gender-neutral registration result that any LFP effect will be concentrated among single mothers specifically, rather than reflecting a general pattern that could explain the broader registration findings.

8.1 Data and outcomes

One main ACS table series is used to study the labor force participation (LFP) channel. It provides a marital-status breakdown of maternal LFP by child-age group, distinguishing between children under 6 and 6 to 17. It is however limited to the LFP rate only, and is thus incapable of distinguishing employment from unemployment. It allows for a refined picture, testing whether LFP effects are concentrated where the constraint is most binding, namely among single mothers. Studying labor force participation rather than employment directly is justifiable here, as many women may be transitioning from outside the labor force to active job search rather than to immediate employment, and a simple employment rate would not capture this margin.

An important data limitation should be acknowledged. The table records children as the unit of observation, and the LFP rate is computed as the share of children in a given arrangement whose mother is in the labor force, rather than mothers directly. Under the assumption of one child per mother in each age-group cell, this is equivalent to the maternal LFP rate. With multiple children per mother in the same age band, the child is double-counted in both the numerator and denominator equally, leaving the rate unbiased but inflating the effective sample size. This table does not allow the distinction between employed and unemployed mothers, so the mechanism test is necessarily limited to labor force participation rather than employment per se.

The identification strategy mirrors the one used for voter registration. The same Callaway-Sant'Anna doubly-robust estimator is applied. Observations are weighted by the time-averaged count of children in each cell, which serves as a proxy for the underlying population at risk. The analysis uses the raw treatment timing definition, with the pre-period binned at $e = -5$ and the post-period balanced at $e = +8$.

8.2 Main results: single versus two-parent households

Figure 8.1 presents the primary result of the mechanism analysis: the 2×2 event-study grid for maternal LFP rates from B23008, with rows corresponding to marital status (single mother versus two-parent family) and columns to child-age group (children under 6 versus children aged 6–17 serving as a placebo).

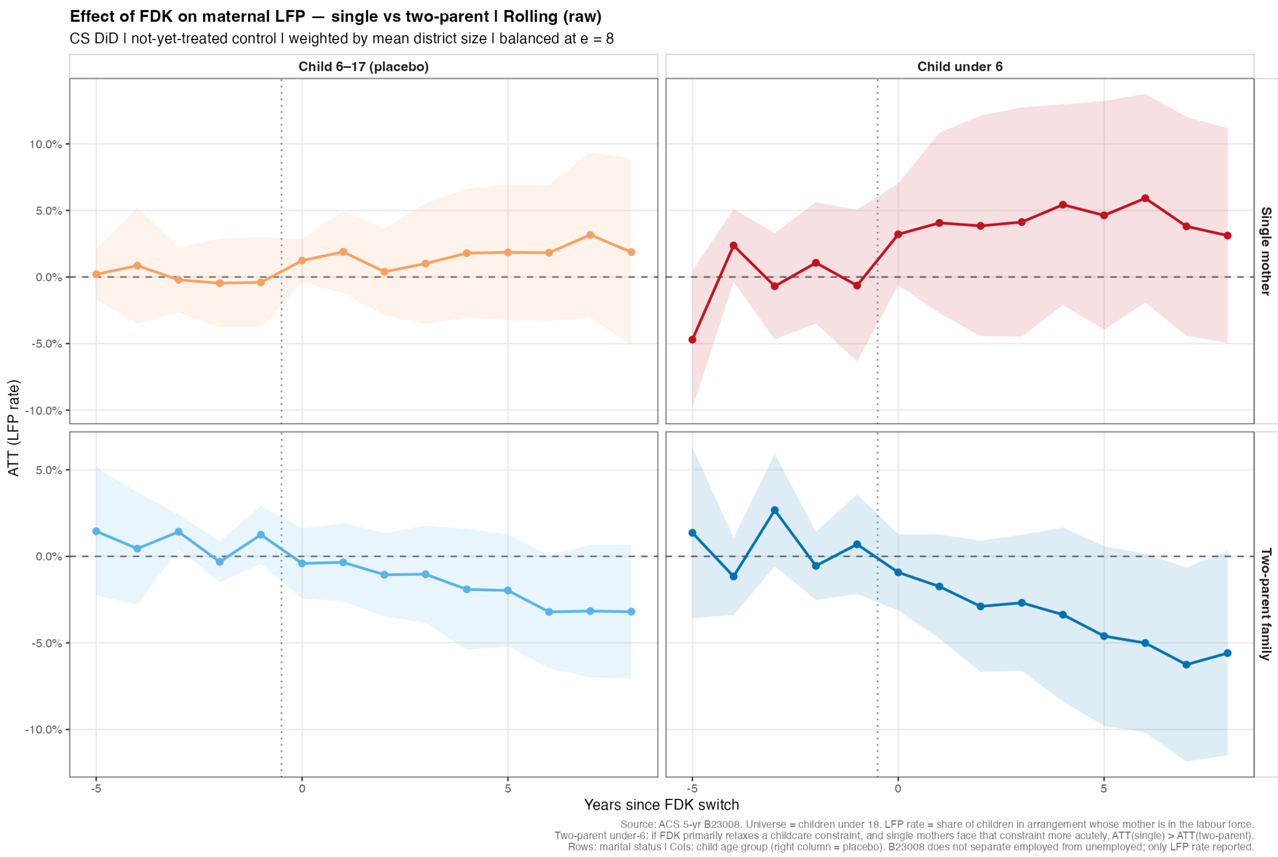

Figure 8.1: Effect of FDK adoption on maternal LFP rate by marital status and child-age group

The four panels of this figure tell a structured story. Beginning with the primary treatment cell, single mothers with children under 6 (top right panel, in red), the event study shows a positive and persistent increase in the LFP rate following FDK adoption. The pre-period estimates are relatively flat, and the post-adoption path jumps at $t = 0$, then rises gradually, reaching roughly 3 to 5 percentage points above the pre-treatment level by $e = +5$ to $e = +8$. This is a substantively meaningful magnitude: a 3 to 5 percentage point increase in labor force participation among single mothers of kindergarten-age children is consistent with what has been found in the literature on childcare expansions more broadly, including the intensive-margin effects documented by Dhuey et al. (2021) using Ontario's rollout. The gradual build-up, following a small immediate jump, suggests that initial gains are present but that the full adjustment takes time, consistent with the idea that labor force entry requires job search and accumulation of work experience rather than an immediate employment transition. It is however not statistically significant.

The bottom right panel, mothers in two-parent families with children under 6, produces a negative and gradually accumulating post-adoption trend, reaching roughly −3 to −6 percentage points by the end of the post-period. Before interpreting this behaviorally, a composition artifact must be considered. Since B23008 records children rather than mothers as the unit of observation, a mechanical compression of the measured LFP rate could arise if the number of children under 6 in two-parent families grows faster in treated districts than in controls post-adoption, inflating the denominator without a proportional change in the numerator. A diagnostic CS DiD on the raw denominator, reported in Figure A.6, finds no evidence of this: the post-adoption ATT on the count of two-parent children under 6 is small and statistically indistinguishable from zero throughout the event window, ruling out the most direct version of this artifact. A more subtle composition concern, whether the distribution of children per mother within the cell is shifting differentially across treated and control districts, cannot be fully tested without PUMS microdata and remains an unresolved caveat.

Taking the result at face value, the most plausible behavioral interpretation is that mothers in two-parent households who were previously working specifically to cover childcare costs exit the labor force or reduce hours once free all-day kindergarten removes that financial pressure. Under this reading the negative effect would represent a revealed preference for non-market time rather than a negative welfare outcome. This interpretation is speculative, however, and the result should be treated as an unresolved puzzle that qualifies rather than supports the broader mechanism claim. The placebo column for mothers of 6–17 year olds shows a mild negative drift consistent in direction with the two-parent treatment cell, suggesting a common underlying trend in these districts rather than a treatment-induced reversal, which further limits the confidence one can place on a causal reading of the negative result in either direction.

In order to validate the single-mother finding, a placebo check is also run for mothers of older children. The single-mother placebo (top left, in orange) shows estimates that hover around zero throughout the observation window with only a slight uptick after a few years. This is what would be expected if the treatment effect identified in the under-6 cell were genuinely attributable to kindergarten access rather than a more general district-level trend, save for the possible presence of kindergarten-age siblings in some households.

8.3 Conclusion on the mechanism with respect to the registration results

The LFP results suggest that a labor market channel is active at the margin where the childcare constraint is most binding: single mothers of kindergarten-age children experience participation gains of 3 to 5 percentage points following FDK adoption, consistent in both sign and magnitude with the broader childcare expansion literature. The gradual build-up in both the LFP and registration event studies is consistent with a multi-step process, since labor market entry and civic integration both take time (Gidengil et al., 2016).

However, the LFP results do not resolve the central tension identified in Section 7.3.1. The registration effect is gender-neutral, with fathers registering at marginally higher rates than mothers following FDK adoption. The LFP effect, by contrast, is concentrated exclusively among single mothers. If the registration gains documented in Chapter 7 were driven primarily by the labor-market-to-civic-engagement channel identified here, one would expect the registration effect to be correspondingly larger for mothers, and largest for single mothers specifically. It is not. These two findings point in different directions and cannot be fully reconciled with the data available. The most defensible interpretation is that the LFP channel captures a real and targeted labor market response among the most constrained households, while the gender-neutral registration effect reflects a broader household-level socialization process that the LFP analysis cannot directly capture. Whether these are two distinct mechanisms or two margins of the same underlying process is a question the current data cannot answer.

Several additional caveats apply. First, the LFP analysis is conducted at the district level using ACS 5-year estimates, which introduces substantial measurement noise relative to the individual-level voter file used for the registration analysis. Second, the inability to distinguish employment from unemployment means that the positive LFP effect observed for single mothers could in principle reflect transitions from outside the labor force to active job search rather than to actual employment. Both would be consistent with increased civic engagement, but the mechanism interpretation would differ. Third, the negative two-parent result remains only partially explained even after the composition diagnostic, and should not be treated as confirmatory evidence for any specific channel. A complementary analysis using individual-level PUMS microdata linking maternal employment status directly to voter registration records would be needed to establish a cleaner causal link between these two outcomes, and represents a clear direction for future research.

9. Robustness, threats to identification, and extensions

9.1 Threats to identification

Before getting into the robustness checks performed for the main results, it is important to clearly explain the specific threats to identification that the design faces, and to then assess properly how serious each of them is likely to be in this setting. The identification strategy relies on the staggered adoption of FDK across Pennsylvania school districts generating quasi-random variation in the timing of treatment, conditional on district and year fixed effects. Four distinct threats to this assumption are worth concretely discussing.

Strategic timing and correlated reforms

The most fundamental threat to identification is that school boards may decide to adopt FDK strategically rather than quasi-randomly, choosing the timing of expansion in response to anticipated electoral or economic conditions that are themselves correlated with the political outcomes studied in this thesis. Two variants of this concern are worth distinguishing.

The first is electoral anticipation. Pennsylvania school board elections occur in odd-numbered years, while the voter registration panel spans all calendar years. If school boards choose to expand FDK in the year or two immediately preceding a school board election, in anticipation of the registration and participation dividends documented in Chapter 7, then the estimated ATT could suffer from reverse causality. Indeed, in that case the positive post-adoption registration effect would partly reflect pre existing electoral dynamics rather than a response to the policy itself. The flat pre trends in the main event study refute this concern, since if an anticipatory effect were present, it could be expected to produce a positive slope in the registration rate starting one to two years before adoption. Nevertheless, it is acknowledged that pre trend tests are not a perfect diagnostic for strategic timing, since the anticipatory registration response could hypothetically be concentrated in the election year itself rather than spread across earlier periods, thus diminishing the causal interpretability of the small jump at $t = 0$. However, figure A.5 does not seem to suggest any correlation between the treat year being odd and the number of districts being treated then. An additional consideration specific to this setting is that school board elections are partisan in Pennsylvania: members do run under official party labels, but these at times can change, and in many districts especially smaller ones they may be elected by name recognition rather than party affiliation. This institutional feature weakens the credit-claiming mechanism through which strategic timing would operate, since it is less obvious that a school board member in a human influenced, rather than partisan, election would benefit electorally from announcing a kindergarten expansion. While this does not eliminate the concern entirely, it does reduce the plausibility of a very close feedback loop between policy timing and partisan electoral incentives at the school board level.

The second main thing to consider concerns correlated reforms. Kindergarten expansion may be systematically accompanied by other district-level policy changes that independently affect voter registration or civic engagement. The most natural candidates are other education spending increases. For instance, if districts that expand FDK also increase spending on other grades or programs, and if this broader spending increase is what drives the registration effect, the FDK expansion would be a spurious marker rather than a cause. This thus remains an acknowledged limitation of the design although further extensions, if for instance one manages to obtain school board level spending data yearly could address this concern, as it would control for this possible endogeneity.

Measurement error in treatment classification

A second threat arises from the fact that treatment is recovered from administrative enrollment data rather than from official policy records. The threshold rule, a jump from below 70% to above 95% FDK enrollment, sustained in all subsequent years, is mechanical and may misclassify the true treatment year for a district, or worse assign a treatment to a district that was not actually treated.

Upward misclassification could occur if a district gradually increased FDK enrollment across several years without a discrete policy decision, with the enrollment crossing the 95% threshold as a result of a sort of organic drift rather than an administrative mandate or clear policy decision. In such cases, the imputed treatment year would post-date the actual policy change, causing the pre-period to contain some post-treatment observations and attenuating the pre-trend test. Downward misclassification on the other hand could occur if a district's enrollment briefly exceeded 95% due to a single-year anomaly (due to factors such as unusually high take-up, boundary changes, or a recording error) before returning to lower levels. The sustained threshold requirement (the district must remain above 95% in all subsequent years) is designed to guard against this however.

A series of validity checks was performed to limit both types of error. Districts with year-on-year enrollment changes exceeding 20 percentage points were flagged for manual review, and where possible cross-referenced against available school board communications and local press archives. The threshold specification is moreover tested as a robustness check in Section 9.2, where an alternative cutoff combination is applied in order to verify that the results are not sensitive to the exact threshold values chosen. Measurement error of this type would generally be expected to attenuate the estimated ATT toward zero, since a noisily measured treatment dummy contains less signal than the true indicator. If this was present, the positive ATTs documented in Chapter 7 could therefore be seen as conservative estimates of the true effect, which may in fact be larger.

Spillovers across bordering districts

A third concern is the possibility of spillover effects across district boundaries. If FDK adoption in one district affects the political behavior of residents in neighboring districts (for instance, through media coverage of the expansion, through cross-boundary kindergarten enrollment, or through general equilibrium effects on local childcare markets) then the never-treated and not-yet-treated control districts may themselves be partially treated, violating the stable unit treatment value assumption (SUTVA) that underlies the DiD estimator.

The direction of any spillover bias is informative. If FDK adoption in a treated district spills over positively to adjacent control districts, the control group's registration rate would be artificially elevated, and the estimated ATT would in turn underestimate the true effect, thus suggesting potentially even larger impacts. If instead families respond to a neighboring district's FDK expansion by migrating into it (the migration concern is discussed below), the control group's population would shrink, potentially compressing the denominator and mechanically inflating the control group's registration rate. Both scenarios produce downward bias in the ATT.

The county-level clustering analysis provides an indirect check on the spillover concern. Analysis of treatment date differences between and within districts finds that the intraclass correlation of treatment year at the county level is approximately 27.9% (see Figure A.7), meaning that roughly 28% of the variation in treatment timing is attributable to county membership. A permutation test is also done, which rejects the hypothesis of purely random within-county timing, indicating that districts in the same county do tend to adopt in closer proximity in time than random chance would suggest. This county level clustering suggests that within-county controls are likely to be partially exposed to the same local policy environment as treated districts, which would tend to attenuate the estimated ATT. It is therefore again a conservative direction of bias. The robustness checks below perform a spillover analysis to address this concern.

Migration

A final threat in this setting is selective migration. If FDK adoption makes a district more attractive to families with young children, who value free all-day kindergarten as a consumption amenity, then districts that expand FDK may attract inward migration of precisely the households most relevant for the registration outcome: parents aged 25–44 with kindergarten-age children. This would inflate both the numerator (new registrations) and the denominator (ACS population 25–44) of the registration rate, but not necessarily in equal proportion. If new migrants register at higher rates than the existing resident population, the ATT would be upward biased; if they register at lower rates (for instance because they are recent arrivals who have not yet established their civic routines in the new location), the ATT would be downward biased.

There are two reasons to think migration bias is limited in this specific setting. First, the scale of intra-state migration in response to school district kindergarten policies is likely small relative to the size of the existing district population, particularly in the short post-adoption window of two to three years that drives most of the ATT estimate. Families do not typically move school districts in direct response to a change in kindergarten provision alone, since moving costs are substantial and kindergarten is only a one-year program. Second, the ACS denominator used throughout is a five-year rolling estimate, which smooths out year-to-year demographic fluctuations and therefore limits the mechanical impact of short-run population movements on the registration rate. An analysis is thus performed, looking at the 5 year registration rates in a CS DiD style regression, with migration as the outcome variable, and it finds no significant pre or post trends (see Figure A.8).

Another main concern in terms of migration is due to the timing of the data collection. The data was downloaded in December 2025, and thus contains the registration information from that time. It does not however contain variables specifying where the person registered to vote initially, and if they changed address. One could assume that once adults become parents they are less likely to move, but this remains a possible threat to identification. This could thus bias the ATT either way due to a misclassification of individuals, which would attribute either too many or too few people to the district. However the orthogonality of migration to the treatment found in the Appendix, Figure A.8, suggests this may not be a systematic threat. It is nonetheless an important possible limitation to keep in mind

9.2 Robustness checks

9.2.1 Alternative estimators

The primary robustness concern in this staggered DiD application is whether the qualitative result depends on the specific assumptions embedded in the Callaway-Sant'Anna estimator. To address this, the main registration result is reproduced under four distinct identification and inference frameworks, presented in Figure 9.1.

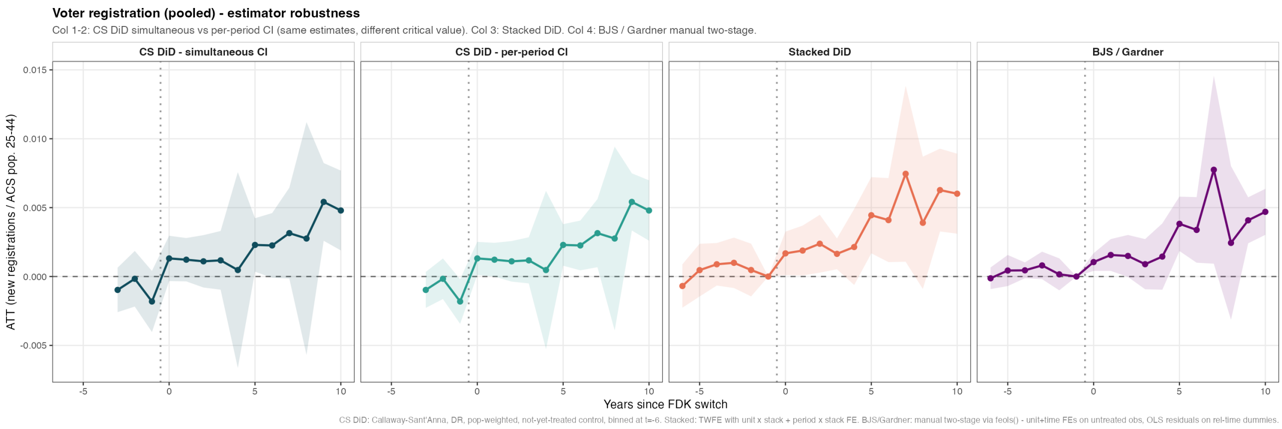

Figure 9.1: Voter registration (pooled) estimator robustness. Each panel presents the same quantity estimated under a different set of assumptions. Col 1–2: CS DiD with simultaneous versus per-period confidence intervals (identical point estimates, different critical values). Col 3: Stacked DiD (Baker et al., 2022). Col 4: BJS/Gardner two-stage imputation (Borusyak et al., 2024).

The four panels tell a consistent story. Across all estimators, the pre-adoption period is flat and close to zero, and the post-adoption path is positive, builds gradually over time, and reaches its highest point in the $t = +8$ to $t = +10$ window. This qualitative agreement is the most important robustness result, since the four estimators differ substantially in their identifying assumptions. The CS DiD estimator relies on semiparametric doubly-robust reweighting, the Stacked DiD avoids negative weighting by restricting each cohort's control group to not-yet-treated districts within a clean sub-panel, but otherwise runs standard TWFE on the stacked data, and the BJS/Gardner two-stage imputation fits unit and time fixed effects on untreated observations only, then recovers the treatment effect as the residual from that imputed counterfactual. The fact that all three substantively distinct procedures produce the same directional finding, with similar point estimates particularly in the $t = +1$ to $t = +6$ range, substantially reduces the concern that the result is an artifact of any particular identifying assumption.

There are however a few quantitative differences. The Stacked DiD point estimates (Column 3) are somewhat larger than the CS DiD in the later post-adoption periods, reaching close to 0.007 at $t = +8$ compared to roughly 0.005 for CS DiD. The BJS/Gardner estimates (Column 4) are noisier, with wider confidence bands, which is consistent with the fact that this estimator uses a smaller effective sample (only untreated observations in the first stage) and does not smooth across cohorts in the same way as CS DiD. Neither of these differences is large enough to change the substantive conclusion, and if anything the Stacked DiD result suggests the CS DiD may be slightly conservative. Interestingly these two specification reach statistically significant results earlier, from $t = 5$ onward, save for $t = 8$ which across all specifications is a very noisy period.

The comparison between Columns 1 and 2 (CS DiD with simultaneous versus per-period confidence intervals) is also informative for inference. The simultaneous bands are wider, as expected given that they control joint coverage over the entire event-study path rather than pointwise coverage at each period. Under the per-period CI, more post-adoption periods are individually statistically significant at the 5% level, while under the simultaneous CI statistical significance is reached primarily at the later horizons ($t = +9$ and $t = +10$). Both are appropriate for different inferential questions: the simultaneous CI answers whether the pre-trend is jointly zero (it is) and whether the post-adoption path as a whole departs from zero (it does, in the later periods), while the per-period CI gives a more granular period-by-period assessment.

Figure 9.2 extends this analysis by adding the three CS DiD specification variants from Chapter 7 (DR weighted never-treated, DR unweighted never-treated, DR weighted not-yet-treated) alongside the alternative estimators, forming a six-panel grid that covers the full range of specifications explored in this thesis.

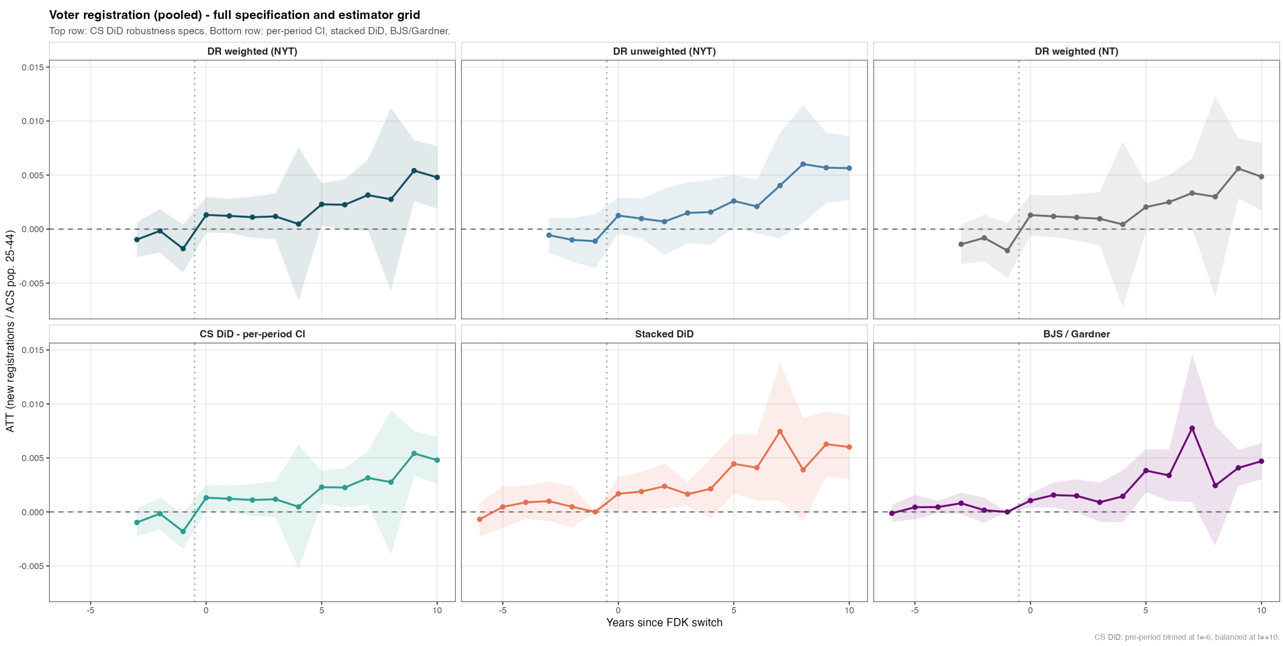

Figure 9.2: Voter registration (pooled). All panels show adults 25–44.

The six-panel grid reinforces the conclusion reached before. Whether the comparison group is the never-treated or not-yet-treated, whether weights are included or not, and whether the estimator is semiparametric or imputation-based, the post-adoption path is positive and the pre-period is flat. The

DR unweighted (NT) and DR weighted (NYT) specifications produce point estimates and event-study shapes very close to the primary DR weighted (NT) specification, confirming that the result is not sensitive to the weighting scheme or the choice of comparison group. The only notable difference across the six panels is in the width of the confidence bands, which is primarily a function of the effective sample size and the inference approach rather than a reflection of genuine instability in the point estimates.

Figure 9.3 presents the gender breakdown under two CI specifications, providing a robustness check on the heterogeneity results discussed in Section 7.3.

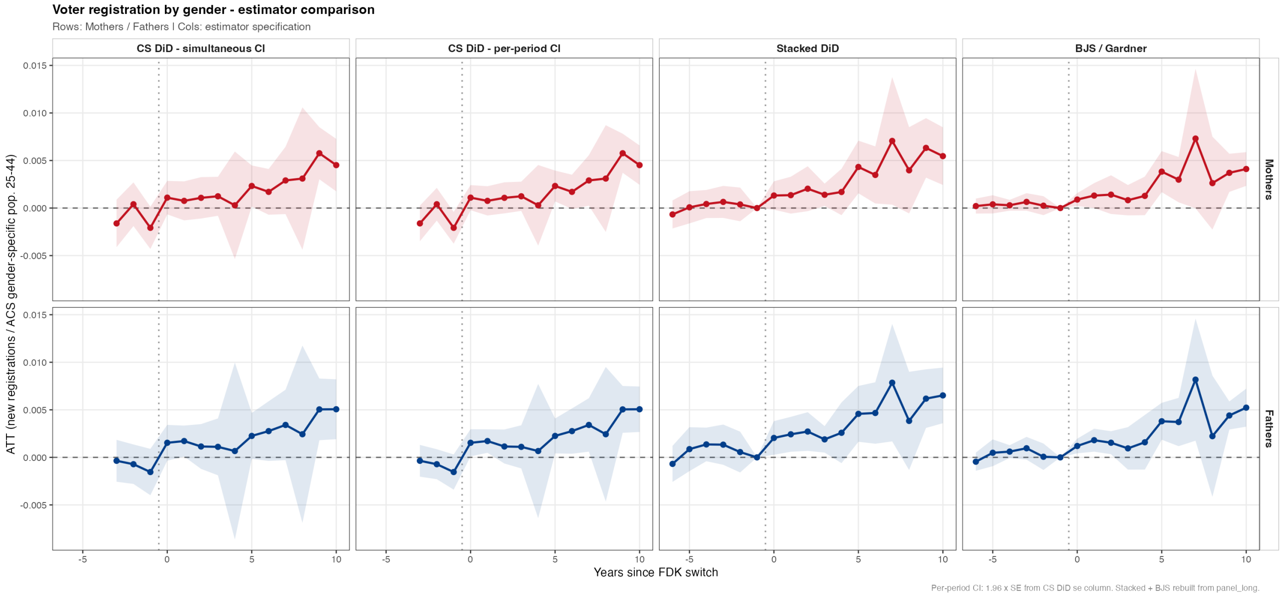

Figure 9.3: Voter registration by gender across specifications

For both parents, the two CI specifications produce nearly identical point estimates and event-study shapes, as expected since the only difference between them is the critical value applied to the same standard errors. The positive and gradual post-adoption build-up is present for both genders under both specifications. Consistent with the discussion in Section 7.3, both series show similar magnitudes, with fathers having only slightly larger point estimates at longer horizons. This pattern holds regardless of whether simultaneous or per-period critical values are used, confirming that the absence of a strong gender concentration in the registration effect is a feature of the data rather than a consequence of the inference approach.

9.2.2 County-level clustering and spillovers

A specific robustness concern raised in Section 9.1 is that treatment timing may be correlated within counties, which would mean that district-level clustering of standard errors understates the true sampling uncertainty if the relevant level of dependence is the county rather than the district. Figure 9.4 provides the key diagnostic for this concern.

The strip chart reveals a mixed picture. On one hand, there is some degree of within-county clustering: in many counties the dots cluster relatively close to the county median (the red diamond), and the earlier analysis reported an ICC of approximately 28%, indicating that roughly a quarter of the variation in

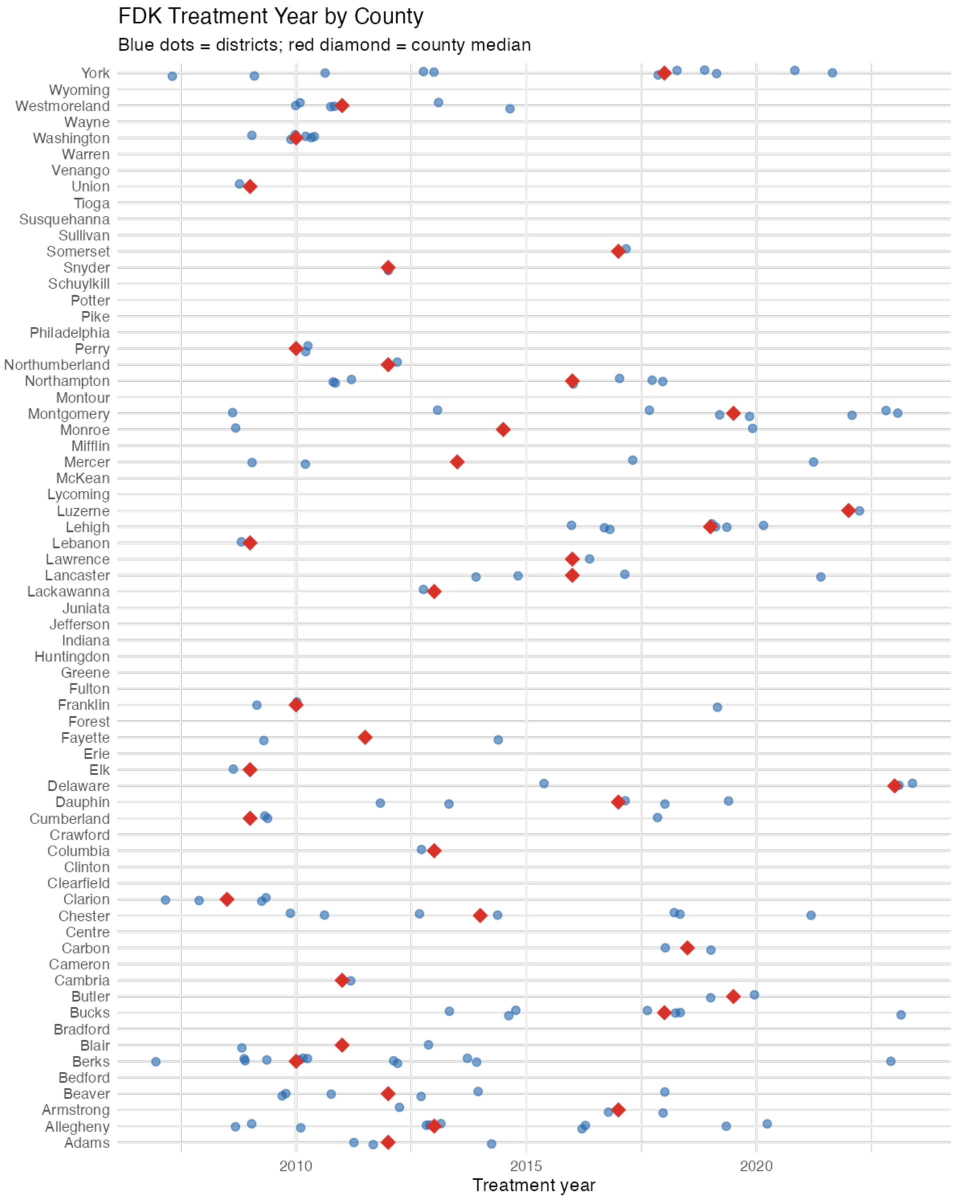

Figure 9.4: FDK treatment year by county

treatment timing is attributable to county membership. On the other hand, the within-county spread is substantial for many of the larger counties. Counties such as York, Berks, Chester, Montgomery, Bucks, and Allegheny each contain multiple districts that adopted FDK in different years, often spanning a decade or more. This within-county variation is precisely what the district-level identification strategy exploits: two districts in the same county, facing the same county-level fiscal environment and political context, nonetheless adopted FDK in very different years, which argues against a pure county-level confound as the driver of the results. Running the main specification while clustering the standard errors at the county level in order to take the largest unit in which there is some degree of correlation yields estimates of the same magnitude. As presented in Table A.1, the ATTs are as significant and robust across all specifications. Considering potential spillovers, an additional robustness is run excluding from the control group districts in counties in which other districts have been treated within 2 years. As shown in Table A.2, the results remain positive but shrink and lose statistical significance. This may be attributable to a series of factors, but the most likely one is the much smaller, potentially distorted, sample size.

9.2.3 Placebo checks and alternative classifications

Two additional robustness checks are relevant for the current analysis. The first is a placebo adoption date test: randomly reassigning treatment years across districts and re-estimating the CS DiD would verify that the observed positive post-adoption ATT is not a spurious artifact of the estimation procedure applied to any time series with the right panel structure. The second is an alternative treatment threshold specification: replacing the 70%/95% thresholds with an alternative that is more likely to indicate a proper jump, here <50% to >95% would verify that the treatment classification is not driving the results through the specific choice of cutoffs. The graph in the Appendix, Figure A.9 suggests the trend is broadly the same changing the threshold, although the estimates get statistically significant faster. The parallel trends are however violated in this case, which suggests there may be differences in districts that never switch and those that see such a large jump. This may be due to a reduced sample size (23 fewer treated districts) or inherent simultaneous political differences. This supports using the more 'gradual' 25% jump specification as the baseline, as it is more comparable with the counterfactual. Nevertheless the general registration trend seems consistent with the previously found results. An arbitrary placebo check on the treat year is also run, taking the treat year -7, and as visible in Figure A.10, this gives the expected null result: no treatment effect and no differential trends between the treated and never treated. This thus suggests that the ATT is interpretable as valid, or at the very least not driven by features of the estimation procedure.

An additional robustness check is the implementation of a narrower-age-band, here 30–39 year olds. This addresses possible dilution from the 25–44 window. As shown in Figures A.11 and A.12, as well as Table A.3, the effect strengthens when focusing on the most plausibly-treated age range. The ATT passes to 0.01069 in the pooled specification, essentially tripling compared to the analysis of the wider range. This thus strengthens the registration effect, and suggests that the more conservative age range considered throughout may be attenuating the strength of the effects.

9.2.4 Specification curve

To assess whether the main voter registration result is sensitive to the specific analytical choices embedded in the baseline specification, a specification curve is constructed over the full cross-product of defensible alternatives: three estimators (Callaway-Sant'Anna doubly-robust, Stacked DiD, and BJS/Gardner two-stage imputation), two control group definitions (not-yet-treated and never-treated), two weighting schemes (weighted by ACS population and unweighted), and three age bands for plausible parents (25–44, 25–34, and 30–44. These age bands are intentionally wider than in the previous subsection, in order to stay relatively conservative). This yields 36 combinations in principle, of which 30 returned stable estimates given available sample sizes and panel support. The results are presented in Figure A.13, with specifications sorted left to right by ATT magnitude and the choice made on each analytical dimension displayed in the bottom panel.

The finding is robust. All 30 estimates are strictly positive and all 30 confidence intervals exclude zero at the 95 percent level, so the qualitative conclusion, that full-day kindergarten adoption increases new voter registration among adults of parenting age, holds without exception across every combination of estimator, control group, weighting scheme, and age window considered. The point estimates range from approximately 0.0015 to 0.0065, a factor of roughly four between the smallest and largest estimate. While the magnitude is not invariant, its sign and significance are. The baseline specification (CS DiD, not-yet-treated, weighted, ages 25–44) sits near the center of this distribution, which provides some reassurance that it is not an unusually favorable configuration.

The bottom panel reveals that the primary driver of variation across specifications is the age band rather than the estimator or control group choice. Specifications using the narrower 25–34 window consistently produce larger point estimates than those using 25–44, which in turn exceed those using 30–44. This pattern is substantively interpretable: parents of kindergarten-age children are on average slightly younger than the full 25–44 window assumes, so the narrower 25–34 band may capture a population more directly affected by the policy, while the 30–44 band dilutes the treatment-relevant population with older adults less likely to have five-year-olds at home. By contrast, the estimator and control group rows in the bottom panel show no systematic ordering, which is consistent with the finding in the main robustness section that CS DiD, Stacked DiD, and BJS/Gardner produce qualitatively similar event-study paths when applied to this panel.

10. Policy discussion

10.1 Interpreting magnitudes

The estimated average treatment effect on the treated of 0.33 percentage points in new voter registrations per year, obtained through the use of modern staggered difference-in-differences techniques, while somewhat modest in absolute terms, translates into a non-trivial number of additional people registering to vote when evaluated at the district level. Taking the mean adult population aged 25–44 across treated districts of approximately 6,500 to 7,600 individuals as a baseline, we get with the ATT that simply expanding kindergarten education leads to roughly 20 to 25 additional new registrations per district per year, or cumulative gains of roughly 200 to 250 registrations per district over the post-adoption decade during which the effect have been found to accumulate. Aggregated across the 142 treated districts, this corresponds to a total participation effect of a few thousand additional registrants. These may not be large numbers in the aggregate, but they could nevertheless be consequential at the margin in close legislative races, of which there have been a fair amount in Pennsylvania in the recent years, as it is notoriously a swing state. This could particularly be the case in smaller House districts where individual-level registration gains can shift vote margins meaningfully. In addition, many of the people who actually registered to vote at some point before 2025 register to vote before the age of 25 (here 59.4% of the sample of possible parents). It is thus plausible that those that had not would not unless there is a major reform which spurs them to vote, an important enough election for them to want to register, or some other factor that convinces them. Thus the fact that statistically significant additional effects are found is an important result in itself.

Additionally, this figure should be treated as a lower bound rather than as a precise point estimate of the true effect. As noted in the analysis and robustness sections, the voter file only captures individuals who had actually registered to vote by December 2025, the date at which the data was downloaded, so those who never registered are absent from the numerator. The ACS denominator aims to correct for this by estimating the size of the full adult population at risk in the districts, but the registration rate used is resultingly mechanically truncated from above. The true new-registration incidence which can be attributed to FDK adoption is therefore likely larger than the 0.33 percentage point ATT suggests, especially since all the tables seem to imply that the effect builds over time. In addition as shown in the previous chapter, the main specification studied is rather conservative, and depending on the estimation strategy and the age range studied, the effect could be much larger.

The relative magnitude is also informative. Pre-treatment yearly registration rates in treated districts range from approximately 0.9% to 1.6% depending on the cohort, which implies that the ATT represents a relative increase of roughly 20 to 35 percent with respect to the baseline. This is a substantively large proportional shift in civic engagement, especially given that the stated aim of this type of policy is educational rather than political. In the labor force participation channel, the estimated gains of 3 to 5 percentage points among single mothers of kindergarten-age children are also sizable, and are more generally consistent with the childcare expansion literature. They nevertheless remain statistically insignificant, portraying the findings as suggestive rather than conclusive. They also reinforce the interpretation that the registration effect, while gradual, operates through a meaningful household-level behavioral change.

The relationship between the registration result and the electoral findings in Section 7.2 deserves explicit reconciliation. The arithmetic is consistent with roughly 230 cumulative additional registrants per district over the post-adoption decade, which in close Pennsylvania House races is a non-trivial number. However, the electoral results in Table 7.2 do not show a corresponding Democratic benefit. Democratic vote share and margins are if anything negative and significant at both the House and Senate level, with the sole exception of Democratic incumbent retention at the Senate level. Three explanations are consistent with this pattern. First, boundary mismatch between school and legislative districts means that the registration gains are diluted and potentially confounded by the political composition of the broader legislative constituency. Second, and more substantively, the registration gains documented in Section 7.1 are non-partisan: as shown in Section 7.3.2, the net Democratic minus Republican registration ATT is negative and statistically indistinguishable from zero, which means that even if newly registered voters turn out at existing-voter rates, no Democratic vote-share advantage would be expected in the aggregate. Third, FDK-adopting districts may be trending Republican over the sample period for reasons unrelated to the policy, introducing compositional confounding in the electoral panel that the legislative-level design cannot fully absorb. A reasonable implication from this finding is that incumbents who expand FDK should not anticipate a short-run partisan electoral dividend from the registration gains this analysis documents, even in a competitive state such as Pennsylvania. The possible exception, a credit-claiming benefit for Democratic incumbents at the Senate level, is too fragile and too dependent on a small and imperfectly matched sample to serve as a reliable basis for electoral strategy.

10.2 Political feasibility and sustainability

The findings in this thesis raise a practical question beyond the causal one: if expanding full-day kindergarten increases civic participation without generating reliable electoral rewards for incumbents in the short run, would the political logic of the policy make it self-sustaining over time, or would it instead remain fragile and contingent on fiscal conditions that may not persist?

Several features of the results point toward a potentially optimistic interpretation in terms of political feasibility. The fact that the registration effect seems to be civic rather than simply partisan, as evidenced by the absence of a statistically significant net advantage for either party in new registrations, irrespective of the incumbency status, suggests that FDK expansion is unlikely to generate strong opposition from one party on purely electoral grounds. Both Democratic and Republican registrations increase following adoption, which suggests that the policy can generate benefits across the political spectrum. This pattern is consistent with the view that FDK functions primarily as a household service rather than a partisan signal, and may therefore be less vulnerable to the kind of polarized rollback that has characterized other redistributive policies in the American context.

However, a few features of the design suggest some degree of fragility. The profile of the registration effect over time is a striking example of this. Statistically meaningful gains do not emerge until approximately nine to ten years after adoption, which is consistent with potential electoral dividend which go to a political horizon well beyond a single school board or legislative term. School board members who implement FDK expansions are thus unlikely to benefit electorally in the short run, which may weaken the incentive to expand in the first place. Importantly the regression on school board election was not run here so this is broadly speculative. This idea however is consistent with the broader political economy literature on public goods visibility: while FDK is salient to parents at the moment of adoption, the registration and participation responses that it generates appear to be mediated through slow-moving behavioral channels, notably labor force participation and community integration, rather than through an immediate credit-claiming response.

It is also important to consider the role of local fiscal capacity in the sustainability of the policy. Pennsylvania school districts are financed primarily through local property taxes, with local sources accounting for approximately 51 percent of total revenues. This reliance on the local tax base suggests that the districts which would adopt FDK during periods of fiscal stress would be more likely to be exposed to roll back of the policy if revenues decline. A district that expands FDK and then faces budget constraints in a subsequent year may face pressure to revert to half-day provision, which would disrupt the gradual behavioral adjustment that the evidence suggests underpins the registration gains. The persistence of the treatment effect, as formalized throughout this thesis, is therefore itself a policy variable: districts that sustain FDK over the medium term are more likely to generate the cumulative registration dividends that the event studies document. This creates an argument for state-level financial guarantees that insulate local FDK provision from year-on-year revenue fluctuations. Importantly though in the analysis of the districts to assign treatment status, no such district was identified in the 2007 to 2025 span of time.

Finally, the indirect evidence on credit claiming from the electoral heterogeneity analysis is worth considering here. The absence of a strong incumbent-party asymmetry in registration gains, and the imprecision of the electoral outcome estimates, suggest that parents do not systematically attribute FDK expansions to the specific party in power. In part, this may reflect the institutional structure of Pennsylvania school boards, where formal ballot labels may exist in some settings, but attribution may still be weak in practice, particularly in smaller districts. This weak attribution mechanism suggests that the political sustainability of FDK does not rest on a tight feedback loop between policy provision and electoral reward at the local level, which could either reduce the risk of rapid reversal by a newly elected board of the opposite party, or alternatively reduce the incentive for any individual board to pursue expansion strategically.

10.3 Equity considerations

The distributional implications of kindergarten expansion are also important to consider, and the evidence in this thesis raises a few equity considerations that are important to acknowledge in addition to the positive causal findings. The most fundamental concern is the distribution of who benefits from FDK adoption, and who bears its costs.

On the benefit side, the LFP analysis in Chapter 8 provides the clearest evidence of distributional heterogeneity. The labor force participation gains from FDK adoption are concentrated among single mothers with children under the age of six, precisely the household type that would previously face the most binding childcare constraint and for whom free all-day kindergarten represents the largest proportional reduction in out-of-pocket costs or opportunity costs. Single mothers in the treated districts experienced non-statistically significant positive LFP increases of 3 to 5 percentage points following FDK adoption, while two-parent families exhibited a negative or null effect. This pattern is consistent with welfare gains from FDK provision that are not evenly spread across parental household types. Families with a secondary earner available to absorb childcare responsibilities were not necessarily as materially constrained before adoption, and may in some cases have adjusted their labor supply downward once the free provision removed the implicit subsidy to employment. The policy thus functions as a targeted intervention for constrained households even when implemented universally.

On the cost side, the local finance structure of Pennsylvania schools raises a more concerning distributional issue. Since over half of school funding derives from local property taxes, the fiscal burden of expanding FDK from half-day to full-day provision falls disproportionately on the property tax base of the district in question. In a state with per-student funding ranging from \$15,800 to \$162,000 across districts, the practical ability to expand FDK is very unequal. Low-wealth districts, which may be precisely those with the highest concentrations of single-parent households and the greatest potential welfare gains from FDK, are also the least able to finance the expansion without external support. This creates a divergence between where the policy is most needed and where it can most easily be implemented.

The balance table in Section 4.3 documents this divergence in a way that is somewhat different from what one might initially expect. The represented always-treated districts are on average the smallest in terms of population aged 25–44 (4,548, compared to 6,726 for those treated early and 6,908 for those treated late) and have the lowest median household income of any group at \$47,921. At first glance this may seem counterintuitive given that universal FDK is a costly service to provide. A plausible explanation, however, is that these districts were previously eligible for state subsidies under Pennsylvania's Basic Education Funding formula, which provides additional support to districts serving families below 184 percent of the federal poverty line. Since the median household income of always-treated districts falls below this threshold, it seems plausible that state redistribution effectively enabled FDK provision in these lower-wealth contexts without requiring a large local tax base. In this light, the always-treated group is not a counterexample to the equity concern but rather an illustration of what targeted fiscal support can hypothetically aim to achieve.

The never-treated districts, by contrast, are the largest (mean population 25–44 of 8,671) and wealthiest (median household income of \$89,269). This is consistent with the hypothesis that higher-wealth communities may have access to a wider supply of private childcare alternatives, which would reduce the effective demand for publicly provided FDK and thus the political pressure to adopt it. The treated districts shown in the table occupy an intermediate position on both dimensions, with early treated districts being on average smaller and lower-resourced than the districts treated later (\$57,732 versus \$67,218 in median household income), which is consistent with the idea that districts facing tighter fiscal constraints moved earlier, as they may be more likely to receive state support. It is also worth noting that the always-treated and never-treated districts differ meaningfully in their partisan composition: always-treated districts have the highest Republican registration share (0.543) and the lowest Democrat share (0.307), while never-treated districts are comparatively more Democrat-leaning (0.359). However a t-test was not run so it is not clear if these are actually statistically different or not. This pattern is somewhat at odds with the literature suggesting Democratic governments are more likely to expand FDK, and suggests that in this specific context local fiscal conditions and state subsidy eligibility may be more important drivers of adoption than partisan composition alone.

These distributional patterns have direct implications for the external validity of the causal estimates. The causal estimates presented in this thesis are most directly relevant to districts of moderate fiscal capacity and may not generalize symmetrically to the always-treated districts, which benefited from a different funding mechanism, nor to the never-treated, where the absence of adoption may reflect either sufficient private alternatives or a fundamentally different demand environment. Estimating the effect of FDK in the subset of low-wealth districts where fiscal constraints have historically prevented adoption is not possible with the current design, but is arguably a policy-relevant question from an equity standpoint, and a clear direction for future research.

11. Conclusion

11.1 Conclusion

This thesis studies the political economy of kindergarten and asks whether it is in a politician's electoral interest to expand full-day kindergarten. To do so, it combined a descriptive, state-level analysis of kindergarten adoption patterns across the continental United States with a causal district-level study of Pennsylvania, exploiting the staggered adoption of full-day kindergarten across school districts within the Callaway and Sant'Anna (2021) difference-in-differences framework.

The first part of the analysis, which documents the evolution of kindergarten enrollment across the United States from the mid-1980s to the present, looks at the extent to which the timing of mandatory attendance law adoption is predicted by pre-existing political or social characteristics. The state-level TWFE estimates reveal a small positive link between kindergarten enrollment intensity and Democratic legislative seat shares: a ten percentage point increase in the share of kindergarten-age children enrolled is associated with a shift of roughly 2.7 to 3.3 additional net Democratic seats in the state Senate, holding state and year fixed effects constant. The corresponding Senate partisan lean estimates are statistically significant at the 5 percent level across multiple rolling average specifications. It is however important to recall that these results are primarily descriptive and should consequently not be over-interpreted. The endogeneity assessment does however provide some insight that these associations seem to not be primarily driven by reverse causality: the logit predictors of adoption timing reveal that pre-period Democratic legislature control and a Democratic governorship are both small and statistically indistinguishable from zero, and pre-adoption enrollment trends track closely across Democratic, Republican, and split legislatures. These findings are explicitly descriptive by construction, and cannot be given a causal interpretation. They are however consistent with the hypothesis that early childhood education can have politically meaningful downstream effects, and motivate the causal design developed in Part 2.

The second, and main, part of the analysis establishes a positive and statistically significant causal effect of full-day kindergarten adoption on new voter registration rates among adults of plausible parenting age in Pennsylvania. The primary CS DiD estimate yields an average treatment effect on the treated of 0.33 percentage points in new registrations per year, significant at the 1 percent level. This represents a relative increase of roughly 20 to 35 percent above the pre-treatment baseline registration rates in treated districts, and should moreover be interpreted as a lower bound on the true new-registration incidence, given the discussion in the threats to identification section. The event-study path is notably gradual: pre-treatment estimates are flat and jointly close to zero, consistent with the parallel trends assumption, while the post-adoption path accumulates slowly, reaching statistical significance only at $t + 9$ and $t + 10$ under the simultaneous confidence bands. This trend argues against a pure salience or credit-claiming channel, and instead seems more consistent with a behavioral mechanism, which materializes progressively as households adjust to the new childcare arrangement.

The heterogeneity analysis reveals that this registration effect is present for both mothers and fathers, with simple ATTs of 0.300 and 0.359 percentage points respectively, and similar event-study profiles for both. This does not straightforwardly support the hypothesis that the effect is concentrated among mothers, which would be the prediction of the pure time-reallocation channel. The gradual post-adoption build-up observed for both genders is however more consistent with a labor market and community socialization mechanism than with an immediate salience response, as discussed further in Chapter 8. The party-level registration decomposition further shows that both Democratic and Republican registrations increase following FDK adoption, with a net Democratic minus Republican ATT which is statistically indistinguishable from zero. This null result on partisan mobilization is itself informative: FDK expansion does not appear to generate a systematically one-sided registration dividend, at least at the aggregate level. This is consistent with the broader pattern found in the electoral outcomes section, where the effects on vote margins and incumbent win probabilities, while directionally negative for Democrats at the House level, are imprecise and should not be interpreted as conclusive evidence of partisan tilting.

The mechanism analysis in Chapter 8 provides some support for the time-constraint channel however, although some important caveats apply. Single mothers with children under the age of six experience labor force participation increases of approximately 3 to 5 percentage points following FDK adoption, consistent in sign and magnitude with the broader childcare expansion literature, and notably with the intensive-margin effects documented by Dhuey et al. (2021). Mothers in two-parent families, by contrast, exhibit a negative and gradually accumulating LFP response, tentatively interpreted as reflecting a reduction in constrained employment rather than a welfare loss, though this result is more speculative than conclusive. The corresponding placebo columns for older child age groups are broadly consistent with these treatment cell results, supporting the interpretation that the LFP effects are actually attributable to the kindergarten channel rather than to more general district-level trends. The gradual build-up observed in both the LFP and the registration event studies is consistent with a multi-step channel, since both labor market entry and civic integration are processes that take time.

The robustness of the main registration result is supported across four distinct estimators: the primary doubly-robust CS DiD, with simultaneous then pointwise confidence intervals, the stacked DiD of Baker et al. (2022), and the two-stage imputation estimator of Borusyak et al. (2024). Conventional TWFE estimations are also shown as a benchmark and, as expected, appear attenuated in this setting. All of the staggered DiD estimators produce qualitatively consistent positive post-adoption paths with flat pre-periods. The comparison between CS DiD and TWFE is moreover illustrative of the identification challenge, as the conventional estimator produces attenuated and in some periods negative post-adoption estimates, consistent with negative-weighting bias under staggered adoption, which further justifies the use of the more modern estimator. Moreover, a specification curve finds significant positive ATTs across all studied specifications.

Overall, the findings suggest that full-day kindergarten expansion has a meaningful and statistically robust effect on civic participation, operating primarily through gradual behavioral adjustment rather than immediate credit-claiming. Whether this translates into direct electoral rewards for incumbents remains an open empirical question, as the sample available for the electoral analysis is not sufficient to detect effects of the magnitude suggested by the registration results.

11.2 Limitations and future research

Several limitations of this thesis deserve to be made explicit, in order to qualify the interpretation of the results and to identify directions for future research.

The most fundamental limitation is the scope of the electoral analysis. The central policy question of this thesis can be most directly answered by studying the electoral outcomes of the decision-making bodies themselves, that is, the school boards responsible for adopting full-day kindergarten in Pennsylvania. This analysis was not implementable here, both because consistent school board composition data is not centrally archived across all 500 Pennsylvania districts and because individual-level factors such as retirements and uncontested seats would make results difficult to generalize, as discussed in Chapter 7. The legislative-level electoral analysis provides a partial substitute, but suffers from the boundary mismatch between school districts and legislative constituencies, and as a consequence the electoral estimates are underpowered. The imprecision of these estimates is not necessarily attributable to an absence of effects however, and may rather just suggest insufficient statistical power. Future work that obtains school board election results at the district level, that focuses on states where school district boundaries map more cleanly onto electoral constituencies, could provide a more direct and intuitive test of the credit-claiming mechanism.

A second limitation is the construction of the treatment variable itself. Treatment is recovered from administrative enrollment data rather than from official policy records, since school board decisions in Pennsylvania are not centrally archived and implementation often occurred gradually. The two-threshold rule applied throughout is designed to identify discrete, persistent policy switches rather than organic drift in enrollment, and a series of validity checks was performed to limit misclassification, as detailed in Section 9.2.3. Nevertheless, measurement error in the imputed treatment year cannot be fully excluded. This type of measurement error would generally be expected to attenuate the estimated ATT toward zero, which suggests that the results presented here are conservative estimates. Future work that obtains complete administrative records of school board votes or official policy announcements would enable the obtention of cleaner and more precisely dated treatment definitions.

A third limitation concerns the mechanism analysis in Chapter 8, which is constrained by the nature of the available ACS data. The LFP table used does not enable the distinction between employed and unemployed mothers to be made, which means that the positive LFP response among single mothers could in part reflect transitions from out of the labor force to active job search, rather than to actual employment. Both are consistent with subsequent civic engagement gains, but the interpretation of the mechanism would differ. More problematically though, the mechanism analysis is conducted at the district level using ACS five-year estimates, which introduces substantially more measurement noise than the individual level voter file used for the registration analysis. Establishing a cleaner causal link between the LFP and registration channels would require individual-level data linking maternal employment status directly to voter registration records, which is not currently feasible in the Pennsylvania administrative context. This represents another potential avenue for further research.

Fourth, the findings of this thesis are derived entirely from the Pennsylvania context, which thus limits their external validity in ways that are worth being explicit about. Pennsylvania is a useful case study precisely because it delegates kindergarten decisions to the LEA level, generating rich within-state variation, and because it is a politically competitive swing state with high-quality administrative data. These features may not generalize to states with different institutional structures, different fiscal environments, or different baseline levels of civic engagement. The five other states that delegate kindergarten policies to local educational agencies (Colorado, Massachusetts, New Hampshire, New Jersey, and New York) would provide natural replication settings, though their very different political economies and demographic compositions would make direct comparison difficult if there is no additional careful harmonization of the empirical design.

Finally, this thesis does not study the cognitive and civic effects of FDK expansion on the children who are actually enrolled in it, nor does it trace longer-run political participation outcomes among those children as they enter adulthood. The literature reviewed in Chapter 2 documents that early education expansions can generate durable effects on civic behaviors through intergenerational channels, and the evidence on intergenerational persistence in political participation (Gidengil et al., 2016) suggests that the registration effects documented here may themselves have downstream consequences well beyond the horizon of the current analysis. Studying these longer-run effects would provide a more complete picture of the political economy of early childhood education, and constitutes a natural extension of this work.

11.3 Policy recommendations

Several policy levers can be identified from the findings and the equity considerations above. These are not exhaustive nor prescriptive per se, but rather should be seen as options whose design and targeting are directly motivated by the empirical results in this thesis.

The first and most direct implication concerns the allocation of state education subsidies. Given that the local finance structure creates an important and systematic divergence between fiscal capacity and demand for full-day kindergarten, state-level redistribution toward low-wealth, high-need districts is a tractable instrument for improving both the reach and the equity of FDK provision. The evidence from the balance table, and in particular the pattern of always-treated districts having the lowest median income while still achieving universal provision, suggests that the Basic Education Funding formula's existing need-based components can be effective when they are sufficiently generous. The evidence in this thesis thus suggests that explicit categorical grants tied to FDK expansion in low-wealth districts could reduce the gap between the never-treated and the rest. Such grants would need to be designed with a sustainability condition, that is, districts should receive funding contingent on maintaining above-threshold FDK enrollment in subsequent years, in order to ensure that the gradual behavioral gains documented in the event studies are not disrupted by fiscal retrenchment. This policy option is of interest if the objectives are of relaxing the time constraint to boost labor force participation, or to increase civic participation more broadly. It should however not be considered as a primarily electoral instrument unless the districts in question are competitive, as the registration gains found here are moderate in magnitude and do not systematically favor any particular incumbent or party.

Secondly, the profile of the registration effect over time has implications for how the political benefits of FDK expansion should be communicated to local decision-makers. If school board members discount political rewards that materialize over a ten-year horizon relative to the immediate costs of property tax increases required to fund expansion, state-level communication strategies that make the long-run civic and electoral benefits salient could help align local incentives with the broader public interest. This is not a purely technical policy instrument, but the evidence on the salience and visibility channel in the credit-claiming literature suggests that how a policy is communicated is itself a determinant of whether its political economy is self-sustaining.

Third, the heterogeneous LFP results by household type suggest that complementary interventions targeted at single-parent households could amplify the welfare and participation gains of FDK. Policies that reduce the remaining barriers to labor force entry for single mothers, such as subsidized after-school care for the hours between FDK dismissal and the end of a standard workday, or childcare assistance for children below kindergarten age, would address the margin at which the FDK-induced relaxation of the time constraint is most binding. The evidence that the LFP effect builds gradually over several years, consistent with a job-search and employment stabilization process, suggests that such complementary supports could hypothetically be most valuable in the first few years after FDK is adopted, as during this period households are still in transition rather than fully adjusted.

Finally, the null result on net partisan registration gains, combined with the absence of a reliable partisan electoral dividend documented in Section 7.2, is consistent with the idea that FDK expansion is realistically unlikely to generate the type of concentrated partisan resistance that could threaten its long-run survival. This is a policy design consideration in itself: a public good whose benefits are perceived as broadly distributed is more durable than one whose distributional consequences are politically contestable. The framing of FDK as a universal household service, rather than as a program targeted at any particular demographic or ideological constituency, is therefore not only empirically accurate given the findings, but also strategically coherent from a political feasibility standpoint.

Bibliography

Baker, Andrew C., David F. Larcker, and Charles C.Y. Wang (2022). “How much should we trust staggered difference-in-differences estimates?” In: *Journal of Financial Economics* 144.2, pp. 370–395.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2024). “Revisiting Event-Study Designs: Robust and Efficient Estimation”. In: *The Review of Economic Studies* 91.6, pp. 3253–3285.

Brilli, Ylenia, Daniela Del Boca, and Chiara D. Pronzato (2016). “Does child care availability play a role in maternal employment and children’s development? Evidence from Italy”. In: *Review of Economics of the Household* 14.1, pp. 27–51.

Bursztyn, Leonardo (2016). “Poverty and the Political Economy of Public Education Spending: Evidence from Brazil”. In: *Journal of the European Economic Association* 14.5, pp. 1101–1128.

Callaway, Brantly and Pedro H. C. Sant’Anna (2021). “Difference-in-Differences with multiple time periods”. In: *Journal of Econometrics* 225.2, pp. 200–230.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein (2010). “The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design”. In: *The Quarterly Journal of Economics* 125.1, pp. 215–261.

Chaisemartin, Clément de and Xavier D’Haultfœuille (2020). “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects”. In: *American Economic Review* 110.9, pp. 2964–2996.

Cohen-Vogel, Lora, J. Sadler, Michael A. Little, Becca Merrill, and F. Chris Curran (2020). “The Adoption of Public Pre-Kindergarten Among the American States: An Event History Analysis”. In: *Educational Policy* 34.7, pp. 1054–1085.

Connelly, Rachel (1992). “The Effect of Child Care Costs on Married Women’s Labor Force Participation”. In: *The Review of Economics and Statistics* 74.1, pp. 83–90.

Cools, Angela (2022). “Parents, Infants, and Voter Turnout: Evidence from the United States”. In: *Quarterly Journal of Political Science* 17.1, pp. 91–119.

Curran, F. Chris (2015). “Expanding Downward: Innovation, Diffusion, and State Policy Adoptions of Universal Preschool”. In: *Education Policy Analysis Archives* 23.79.

Dhuey, Elizabeth, Jessie Lamontagne, and Tingting Zhang (2021). “Full-Day Kindergarten: Effects on Maternal Labor Supply”. In: *Education Finance and Policy* 16.4, pp. 533–557.

Dynes, Adam M. and John B. Holbein (2020). “Noisy Retrospection: The Effect of Party Control on Policy Outcomes”. In: *American Political Science Review* 114.1, pp. 237–257.

Finan, Frederico and Maurizio Mazzocco (2021). “Electoral Incentives and the Allocation of Public Funds”. In: *Journal of the European Economic Association* 19.5, pp. 2467–2512.

Fischel, William A. (2001). *The Homevoter Hypothesis: How Home Values Influence Local Government Taxation, School Finance, and Land-Use Policies*. Cambridge, MA: Harvard University Press.

Fischer, Adrienne, Gerardo Silva-Padron, Lauren Peisach, Matt Weyer, and Carlos Jamieson (June 2023). *50-State Comparison: State K-3 Policies*. Education Commission of the States. URL: https://www.ecs.org/50-state-comparison-state-k-3-policies-2023/.

Gardner, John (2022). “Two-Stage Differences in Differences”. Working paper. Available at https://arxiv.org/abs/2207.05943.

Gibbs, Chloe, Jocelyn S. Wikle, and Riley Wilson (June 2025). *A Matter of Time? Measuring Effects of Public Schooling Expansions on Families*. NBER Working Paper 33948. National Bureau of Economic Research.

Gidengil, Elisabeth, Hanna Maria Wass, and Maria Susanna Valaste (2016). "Political Socialization and Voting: The Parent–Child Link in Turnout". In: *Political Research Quarterly* 69.2, pp. 373–383.

Gottlieb, Jessica and Katrina Kosec (2019). "The Countervailing Effects of Competition on Public Goods Provision: When Bargaining Inefficiencies Lead to Bad Outcomes". In: *American Political Science Review* 113.1, pp. 88–107.

Hoxby, Caroline M. (2000). "Does Competition Among Public Schools Benefit Students and Taxpayers?" In: *American Economic Review* 90.5, pp. 1209–1238.

Jackson, C. Kirabo, Rucker C. Johnson, and Claudia Persico (2016). "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms". In: *The Quarterly Journal of Economics* 131.1, pp. 157–218.

Kimmel, Jean (1998). "Child Care Costs as a Barrier to Employment for Single and Married Mothers". In: *The Review of Economics and Statistics* 80.2, pp. 287–299.

Mani, Anandi and Sharun Mukand (2007). "Democracy, visibility and public good provision". In: *Journal of Development Economics* 83.2, pp. 506–529.

Marchese, Alda, Paola Profeta, and Giulia Savio (2025). "Who cares about childcare? Covid-19 and gender differences in local public spending". In: *International Tax and Public Finance* 32.6, pp. 1990–2010.

Sun, Liyang and Sarah Abraham (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". In: *Journal of Econometrics* 225.2, pp. 175–199.

Tebaldi, Raquel, Franziska Gassmann, and Bruno Martorano (2024). "Re-election incentives and the adoption of early childhood development programmes: The case of the "Happy Child" Programme in Brazil". In: *European Journal of Political Economy*. In press (corrected proof), available online 24 October 2024, p. 102616.

Tebaldi, Raquel and Alysson Lorenzon Portella (2025). "Electoral Rules and the Provision of Early Childhood Education and Care: Evidence from Brazil". In: *Journal of Development Studies* 61.7, pp. 1126–1147.

Teney, Céline, Stephan Dochow-Sondershaus, and Forrest Lovette (2024). "The Gendered Effect of Parenthood on Voting Behaviour in the 2021 German Federal Election". In: *German Politics* 33.1, pp. 22–45.

Yang, Hyunwoo (2023). "What Factors Drive School Districts to Adopt Universal Pre-Kindergarten (UPK): Evidence from Wisconsin". In: *Early Education and Development* 34.5, pp. 1124–1143.

A. Tables and maps

A.1 Maps

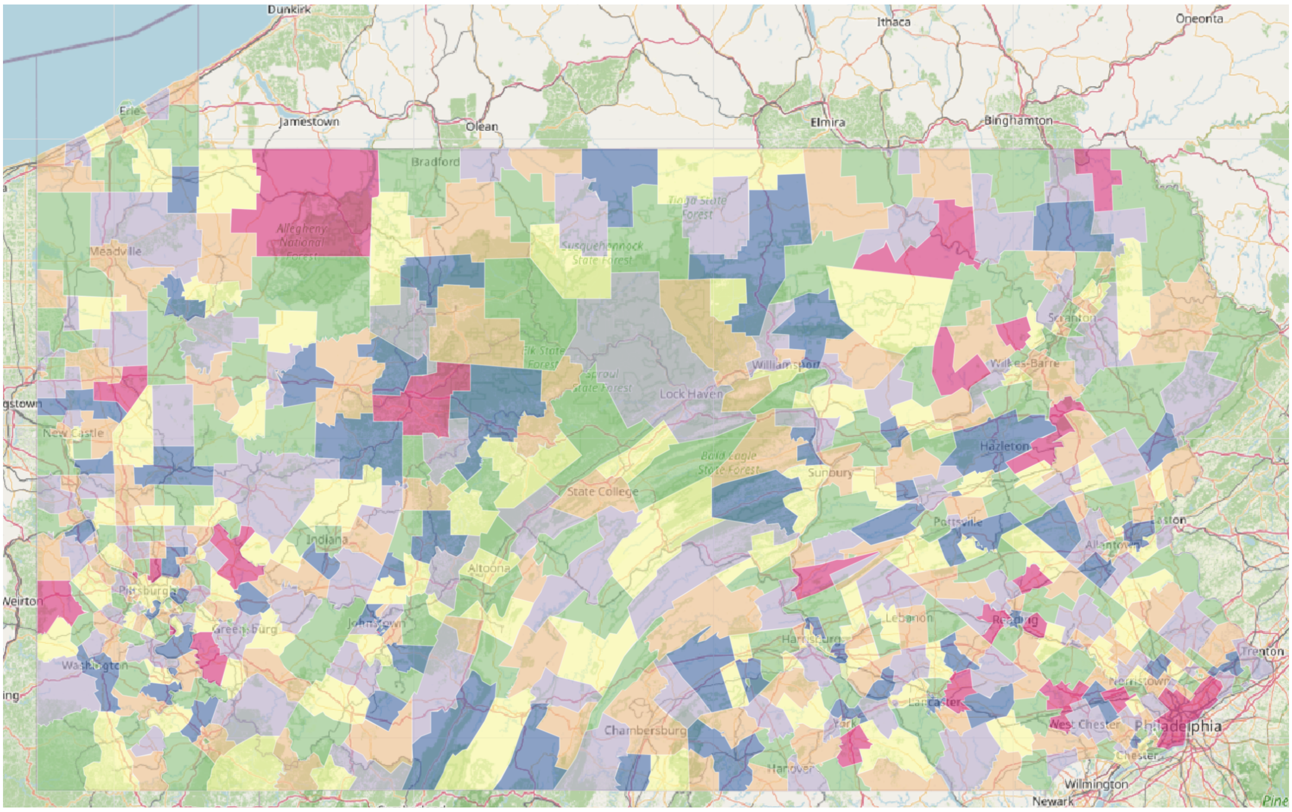

Figure A.1: Map of Pennsylvania school districts

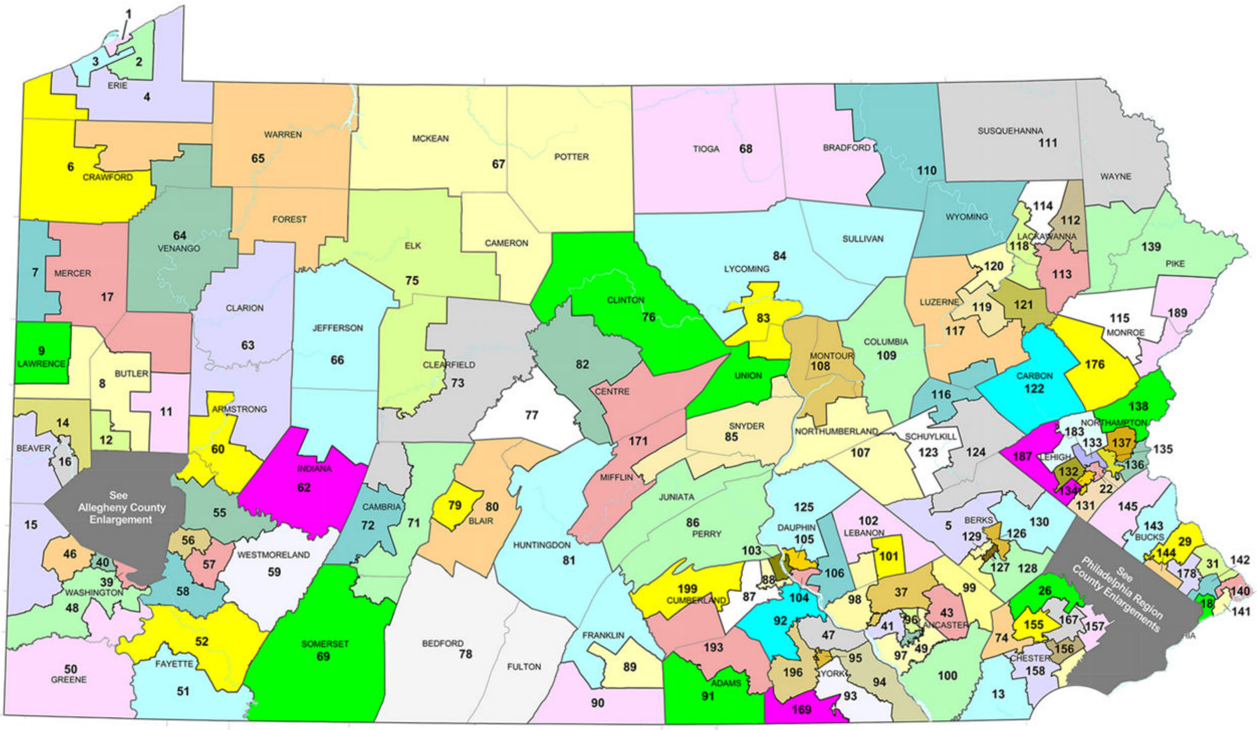

Figure A.2: Map of Pennsylvania state house districts

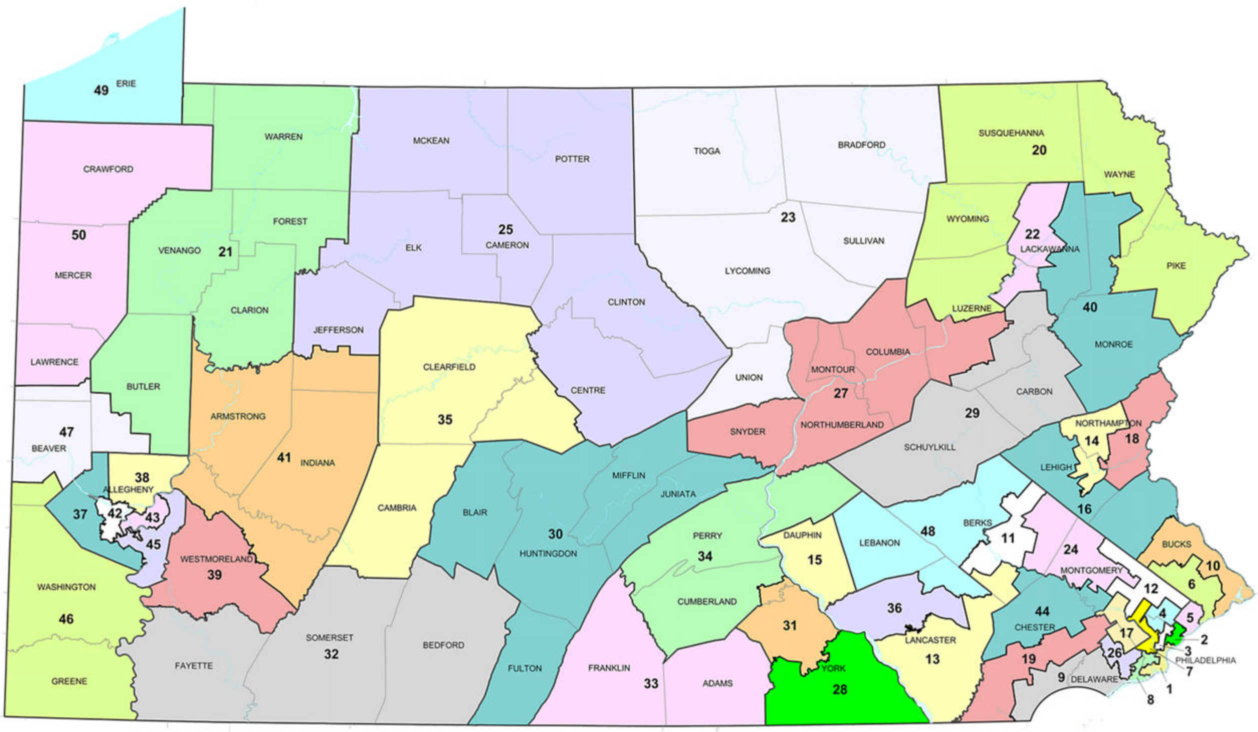

Figure A.3: Map of Pennsylvania state senate districts

A.2 Extra content and checks

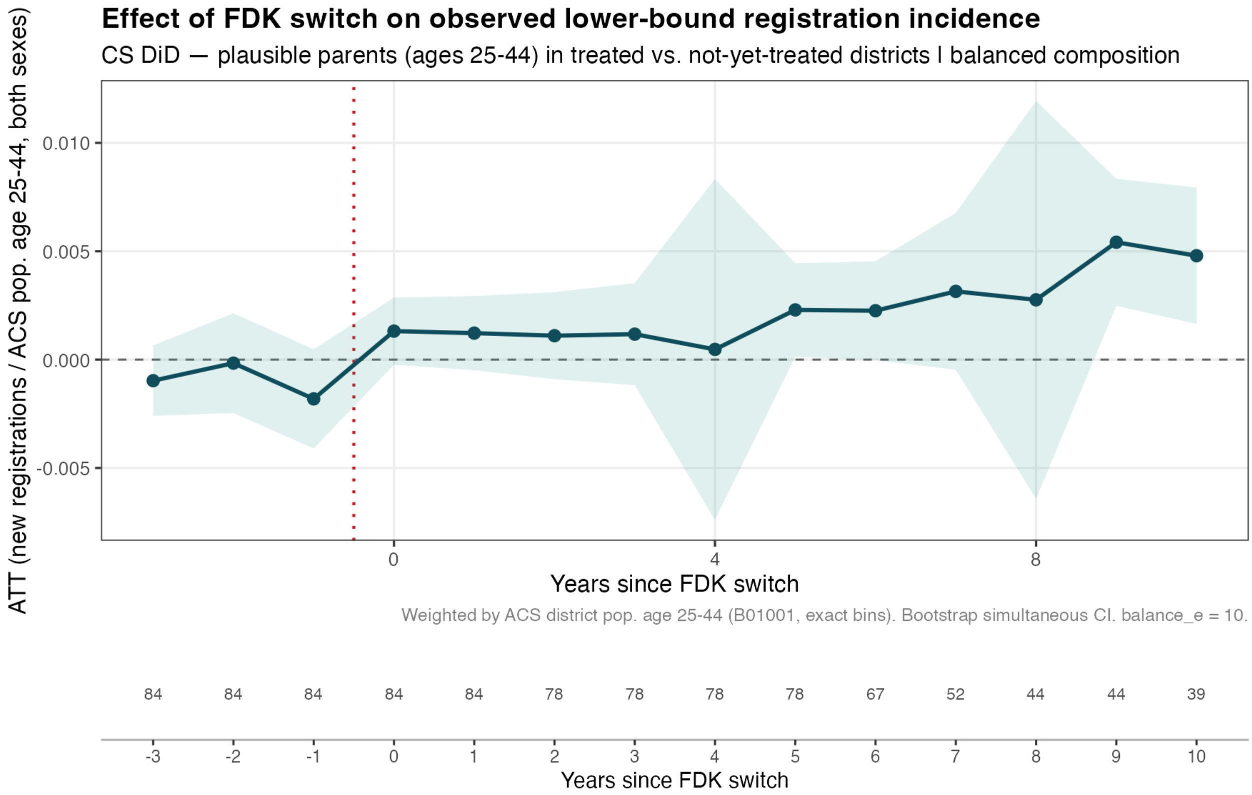

Figure A.4: Parent CS DiD with control group count

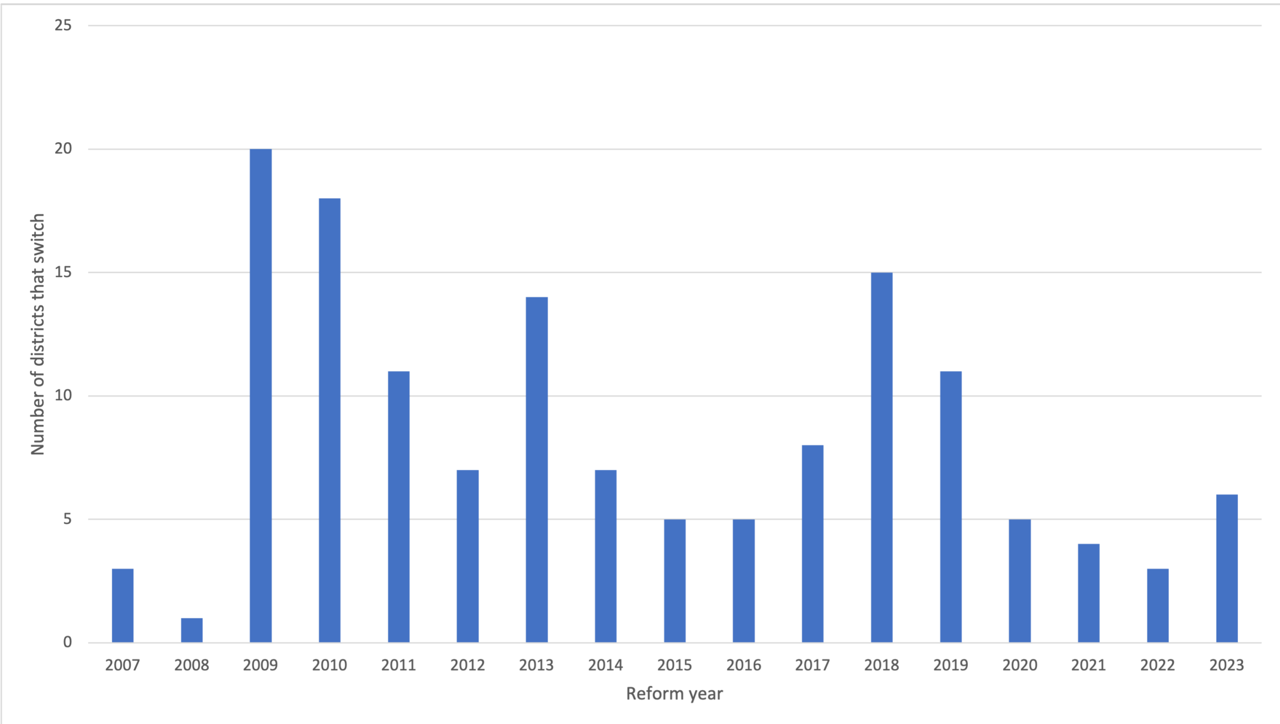

Figure A.5: Number of districts that switch per year

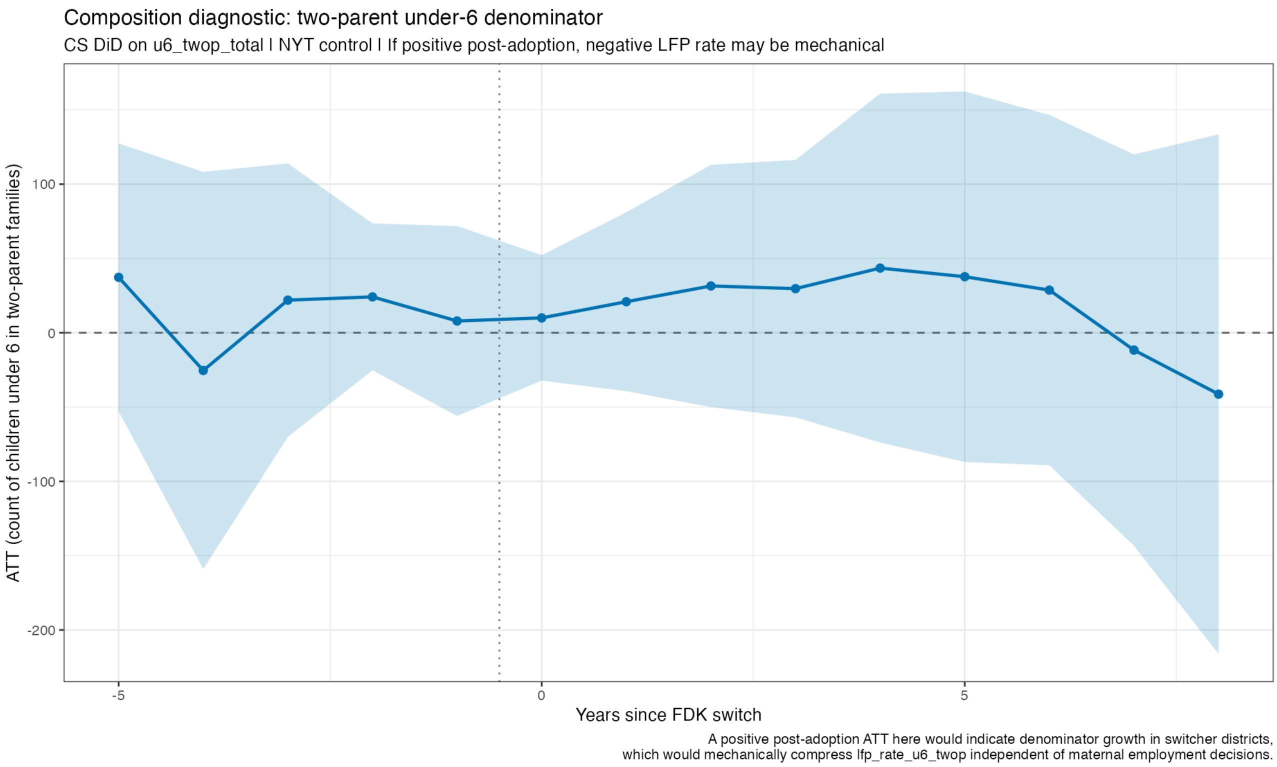

Figure A.6: Diagnostic of the two parent denominator

Statistic	Value
Same-county mean gap (years)	4.18
Cross-county mean gap (years)	5.00
Same-county median gap (years)	3
Cross-county median gap (years)	4
Wilcoxon rank-sum p-value	0.7472
ICC (% variance explained by county)	27.9%
Within-county SD (σ_w, years)	3.77
Between-county SD (σ_b, years)	2.34

Figure A.7: Within- vs. cross-county treatment timing statistics.

Table A.1: Effect of FDK adoption on voter registration — all participation specifications

Specification	ATT (SE)		95% CI lo	95% CI hi
Pooled — DR weighted, not-yet-treated	0.00330***	(0.00115)	0.00104	0.00556
Pooled — DR unweighted, not-yet-treated	0.00327**	(0.00139)	0.00056	0.00599
Pooled — Regression adj., not-yet-treated	0.00330***	(0.00115)	0.00105	0.00555
Pooled — DR weighted, never-treated	0.00337***	(0.00116)	0.00109	0.00565

Note:
*** p<0.01, ** p<0.05, * p<0.10. CS DiD DR estimator unless noted. Outcome: new registrations / ACS pop. 25-44. Not-yet-treated districts as control throughout. Bootstrap simultaneous confidence intervals. Clustering at the county level

Table A.2: Spillover Robustness: Callaway–Sant'Anna ATT Estimates

Subgroup	Specification	ATT	SE	95% CI	N_{ctrl}
Pooled	Baseline	0.00330**	(0.00129)	[0.001, 0.006]	51
	Spillover-robust	0.00096	(0.00092)	[−0.001, 0.003]	14
Mothers	Baseline	0.00300**	(0.00121)	[0.001, 0.005]	51
	Spillover-robust	0.00077	(0.00093)	[−0.001, 0.003]	14
Father	Baseline	0.00359***	(0.00131)	[0.001, 0.006]	51
	Spillover-robust	0.00118	(0.00098)	[−0.001, 0.003]	14

Notes: Callaway–Sant'Anna DiD, DR estimator, not-yet-treated control group, weighted by ACS population aged 25–44 (B01001, exact bins). *Spillover-robust*: control district–year observations are dropped when the control district shares a county with any switcher and the calendar year falls within ±2 years of that switcher's treatment date. N_{ctrl} = number of not-yet-treated control districts. Bootstrap simultaneous CIs (1000 iterations). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ (two-sided, bootstrap critical value).

Placebo and threshold changes

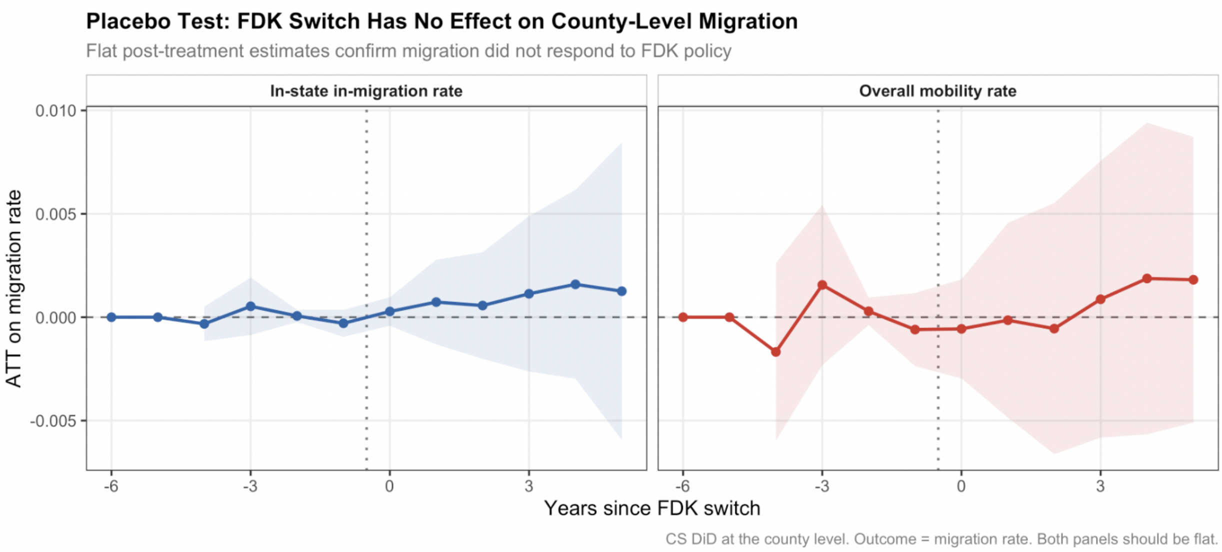

Figure A.8: Placebo check on migration

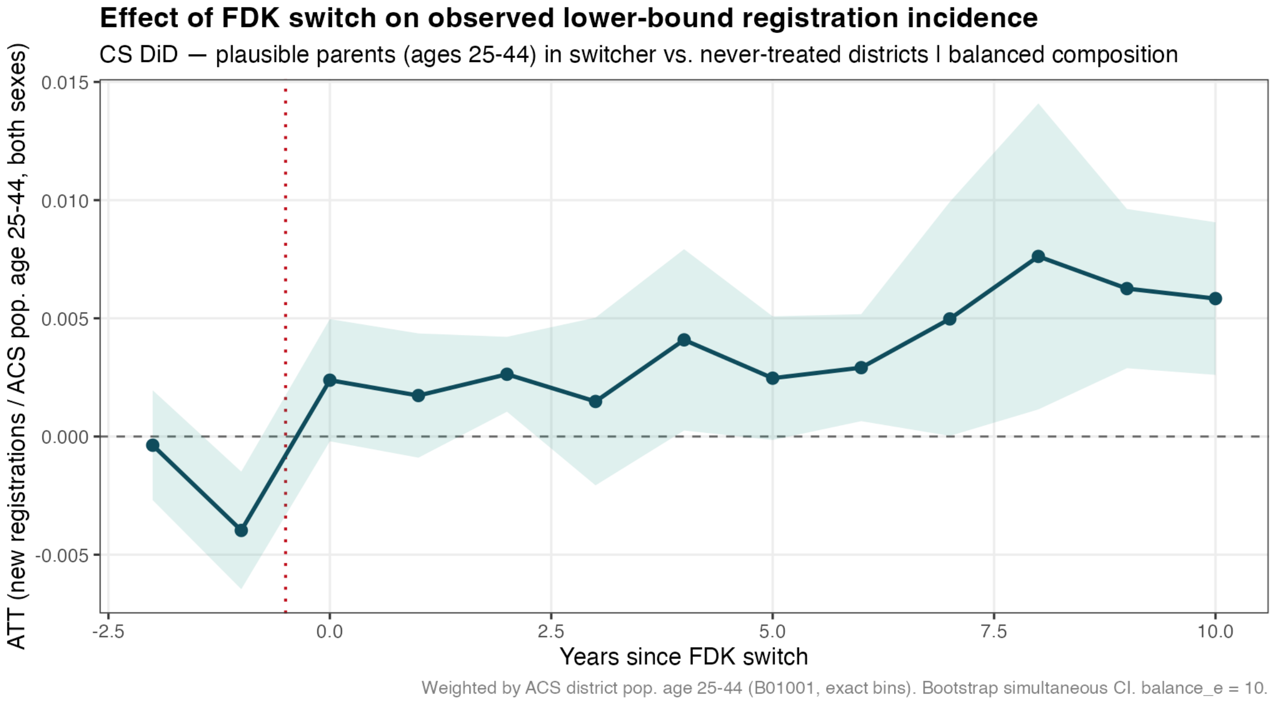

Figure A.9: Change of threshold check, CS DID

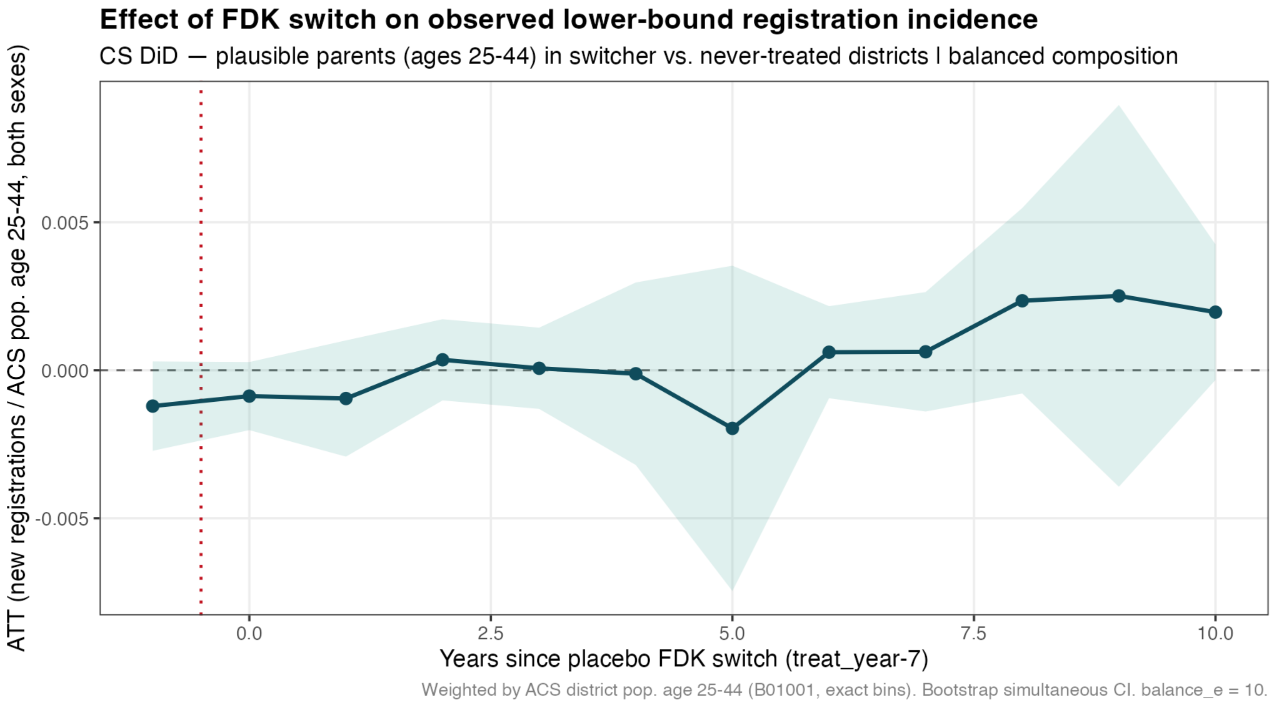

Figure A.10: Placebo treat year, CS DID

Robustness on the age range

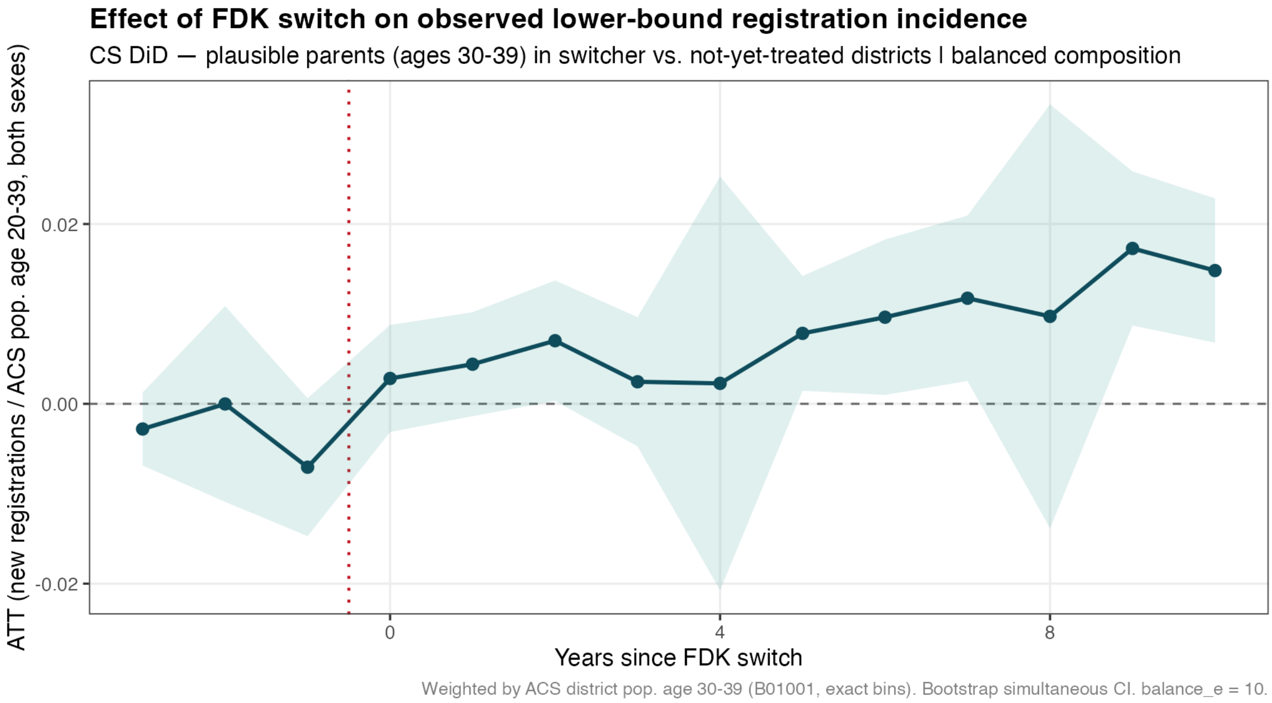

Figure A.11: Main specification 30-39 range

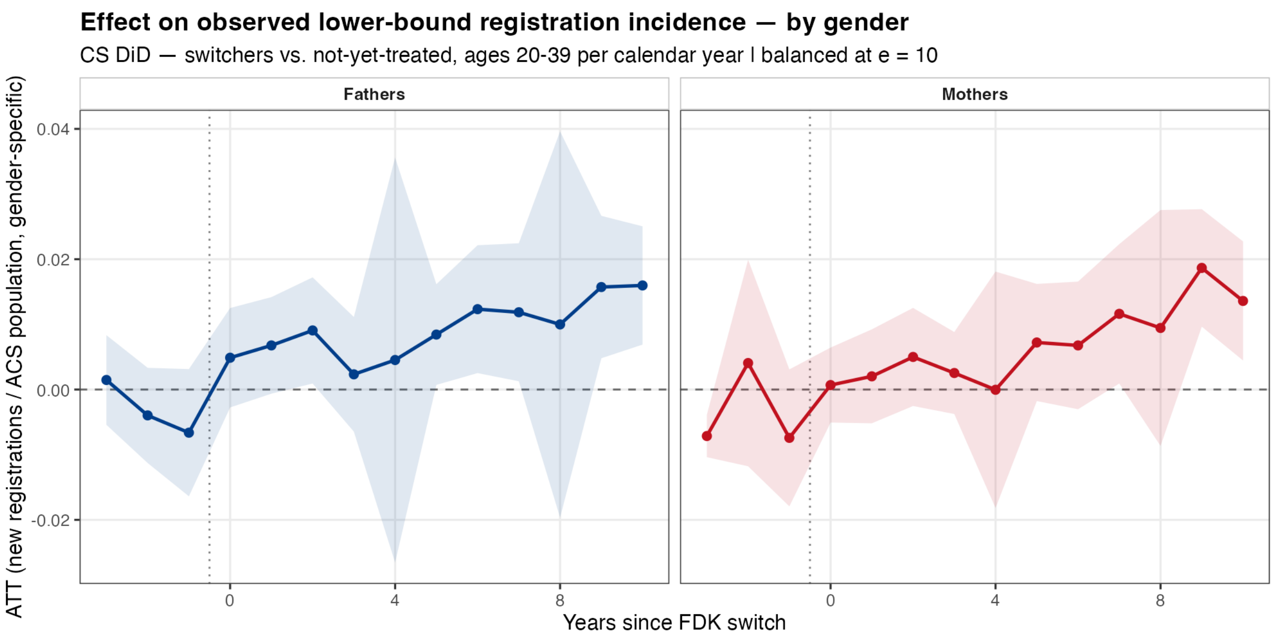

Figure A.12: Main specification by gender 30-39 range

Table A.3: CS DiD estimates: FDK adoption on voter registration

Outcome	ATT (SE)
Registration rate (pooled)	0.01069** (0.00442)
Registration rate (mothers)	0.00981** (0.00401)
Registration rate (fathers)	0.01151** (0.00467)
Democrat registrations	0.00343*** (0.00107)
Republican registrations	0.00427 (0.00306)
Net Dem minus Rep registrations	-0.00084 (0.00257)

Note:

*** p<0.01, ** p<0.05, * p<0.10. Standard errors in parentheses. CS DiD DR estimator, not-yet-treated control. Weighted by ACS pop. 30-39. Positive = more Democratic.

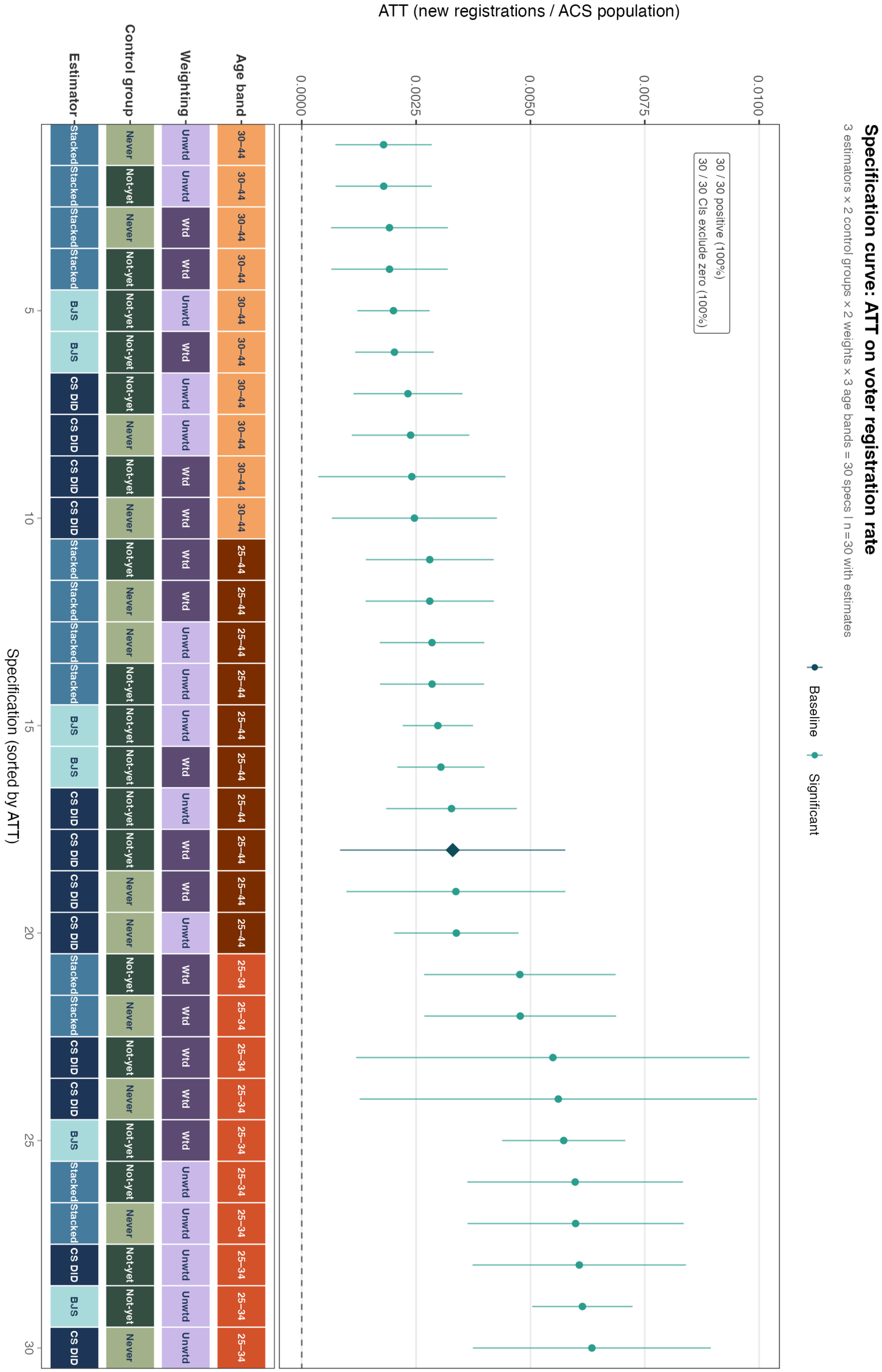

Figure A.13: All possible specifications

B. Additional data details

This appendix documents all data sources used in the thesis. For each source the table reports: the geographic and temporal coverage, the specific files or tables extracted from each source, the variables obtained after cleaning, and the analyses or outputs in which each source appears.

Source 1: Pennsylvania Department of Education — LEA Enrollment Files

Attribute	Detail
Access	Annual Excel files downloaded from the PA Department of Education public data portal (https://www.education.pa.gov).
Unit of observation	Local Educational Agency (LEA) × school year.
Geographic coverage	All ≈667 LEAs in Pennsylvania (500 school districts + charter / vocational LEAs).
Temporal coverage	2008–09 through 2024–25 school years
Raw variables	
Identifiers	AUN (Administrative Unit Number), LEA Name, LEA Type, County.
Enrollment counts	Full-day kindergarten enrollment (`kf_total`), half-day kindergarten enrollment (`kh_total`), combined kindergarten total (`k5_total`), Grade 1 enrollment (`x1`), Grade 12 enrollment (`x12`), enrollment in all other grades.
Derived variables	
Full-day share	full-day KG / total KG enrollment (proportion).
All-in indicator	binary flag equal to 1 when the district reports no half-day enrollment and full-day $\geq 95\,\%$.
Treatment status	`treat_year`: first year satisfying the two-threshold switch rule, `switcher` / `always_treated` / `never_treated` indicators.
Event time	`event_time` = *year* − `treat_year`.
Used in	Treatment classification algorithm; County treatment timing; all CS DiD and TWFE analyses in Chapters 7–9; pre-treatment balance table; Alternative threshold robustness check.

Source 2: Pennsylvania Board of Elections — Statewide Voter File (SURE)

Attribute	Detail
Access	Tab-delimited flat files (`Statewide/*FVE.txt`) and election-map files (`Statewide/*Election Map*.txt`), downloaded December 2025
Unit of observation	Individual registered voter (8 911 681 records).
Geographic coverage	All 67 Pennsylvania counties.
Temporal coverage	Registration records from any date up to December 2025; election-participation history covering up to 40 elections per voter, primarily 2006–2025.
Raw variables (voter spine)	
Identifiers	`VoterID` (unique SURE ID), `County`.
Demographics	`Gender` (M/F), `DOB` (date of birth).
Registration	`RegistrationDate`, `Status`, `StatusChangeDate`, `Party` (registered party affiliation).
Geography	`city`, `state`, `zip`, `PrecinctCode`, `PrecinctSplit`, `SchoolDistrict`, `StateHouse` (PA House district), `StateSenate` (PA Senate district).
Participation	`LastVoteDate`.
Raw variables (election history)	
Per-election records	Up to 40 election slots per voter: `VoteMethod` (whether and how the voter cast a ballot) and `ElectionParty`; linked to `ElectionName`, `ElectionYear`, `ElectionType`, and `ElectionDate` via the county-level election map.
Derived variables	
Age variables	`BirthYear`, `BirthMonth`, `BirthDay`; `KG_year` (imputed kindergarten cohort year, accounting for the September 1 cutoff).
Registration timing	`RegistrationAge` (age at registration); `RegYear` (calendar year of first registration).
Outcome: registration rate	`reg_rate` = new registrations among 25–44-year-olds / ACS population 25–44; computed separately for all adults, mothers (female), and fathers (male); and by party
Partisan majority proxy	`district_majority`: most-represented registered party in a district prior to the reform year, used as a proxy for incumbency.

Attribute	Detail
Legislative-district mapping	Each voter's school district is linked to the most common PA House and Senate district; treatment is then aggregated to the legislative level.
Used in	Primary outcome series (Chapters 7–9): registration rate (pooled, by gender, by party); electoral outcomes; heterogeneity by incumbency; balance table ; migration placebo .

Source 3: American Community Survey (5-year) — Table B01001, Sex by Age

Attribute	Detail
Access	Census API via `tidycensus` and direct HTTP calls
Unit of observation	School district × ACS year.
Geographic coverage	All PA unified school districts.
Temporal coverage	ACS 5-year vintages 2010–2024.
Variables extracted	
Male 25–44	`B01001_011E` (25–29), `B01001_012E` (30–34), `B01001_013E` (35–39), `B01001_014E` (40–44).
Female 25–44	`B01001_035E` (25–29), `B01001_036E` (30–34), `B01001_037E` (35–39), `B01001_038E` (40–44).
Derived variables	Male, female, and total populations as a sum of bins
Used in	Registration rate denominator; gender-specific event-study denominators ; CS DiD population weights; balance table .

Source 4: American Community Survey (5-year) — Table B19013, Median Household Income

Attribute	Detail
Access	Census API in `5_balance_table.R`; variable `B19013_001E`; same geography and vintage as Source 3.
Unit of observation	School district × ACS vintage year.

Attribute	Detail
Temporal coverage	ACS 5-year vintages 2010–2024.
Variables extracted	median household income in dollars (negative values recoded to missing).
Used in	Pre-treatment balance table, column *Median household income*; discussion of fiscal heterogeneity.

Source 5: American Community Survey (5-year) — Table B23008, Employment Status of Women in Families by Child Age and Marital Status

Attribute	Detail
Access	Census API via `tidycensus` in `2a_c_LFP_marital.R`.
Unit of observation	State (or district) × ACS vintage year.
Temporal coverage	ACS 5-year vintages 2010–2024.
Universe note	*Children* (not mothers) are the unit of observation in B23008; the resulting LFP rate equals the share of children in each arrangement whose mother is in the labor force (unbiased under one child per mother; inflated *N* with multiple same-age children). Does *not* distinguish employed from unemployed.
Variables extracted	
Under-6 / two-parent	`B23008_003` (total), `B23008_004` (both parents in LF ⇒ mother in LF), `B23008_006` (mother only in LF).
Under-6 / single mother	`B23008_012` (total with mother), `B23008_013` (mother in LF), `B23008_014` (mother not in LF).
Aged 6–17 / two-parent	`B23008_016`–`B23008_020` (parallel structure to above; used as placebo group).
Aged 6–17 / single mother	`B23008_025`–`B23008_027` (placebo group).
Derived variables	`lfp_rate_single_u6`: maternal LFP rate, single mothers, children under 6 (primary treatment cell). `lfp_rate_twopar_u6`: two-parent families, children under 6. `lfp_rate_single_617` and `lfp_rate_twopar_617`: placebo cells (children aged 6–17).
Used in	LFP mechanism analysis; conclusion on the LFP channel; policy discussion.

Source 6: U.S. Decennial Census / Census Bureau Archives — State-Level Age Population, 1980–2010

Attribute	Detail
Access	Online from the archives "Population Estimates for the U.S., Regions, Divisions, and States by 5-year Age Groups and Sex".
Unit of observation	State × year.
Geographic coverage	All 50 states + D.C.
Temporal coverage	1980–2010 (annual intercensal estimates).
Variables extracted	`Under 5 years`, `5 to 9 years` (5-year age bins); exact single-year age counts for 2000–2009 (columns `Age_0` through `Age_17`).
Derived variables	`imputed_kg_age`: estimated count of kindergarten-age children (≈ `Age_5` from exact counts when available; otherwise `5 to 9 years` / 5). Used as denominator in the kindergarten enrollment share.
Used in	State-level exposure intensity; enrollment around mandatory law enactment; TWFE correlates regressions.

Source 7: American Community Survey — State-Level School Enrollment Tables, 2010–2024

Attribute	Detail
Access	Downloaded CSV files
Variables extracted	Age population: `Under 5 years`, `5 to 9 years`. Enrollment: `Kindergarten Total`, `Kindergarten Public`, `Kindergarten Private`.
Derived variables	`pct_in_kg` (post-2010 segment of the state-year exposure series); population-weighted and unweighted national averages.
Used in	State-level exposure series; TWFE correlates of political outcomes.

Source 8: NCES Elementary/Secondary Information System (ELSi) Generator —
State Kindergarten Enrollment, 1986–2010

Attribute	Detail
Access	File downloaded from the NCES ELSi custom-report generator (`https://nces.ed.gov/ccd/elsi/`)
Unit of observation	State $\times$ academic year.
Geographic coverage	All 50 states + D.C.
Temporal coverage	Academic years 1986–87 through 2009–10 (mapped to calendar year = start year + 1).
Variables extracted	`Kindergarten Students` (total KG enrollment, including public and private).
Derived variables	Used as the numerator of `pct_in_kg` for the pre-2010 period; merged with Census age-population denominators to form the continuous 1986–2024 exposure series.
Used in	State-level exposure series; TWFE correlates

Source 9: Kindergarten Policy Dates — Manually Compiled Cross-Walk

Attribute	Detail
Access	Excel file `Kgt policies dates.xlsx`, compiled manually from the Education Commission of the States 50-state comparison (Fischer et al., 2023) and supplementary state legislative records
Unit of observation	State (cross-sectional).
Geographic coverage	All 50 states + D.C.
Variables	`State`; `Pupil Attendance in Kindergarten` (mandatory M / not mandatory); `Year Mandatory Attendance Law Enacted`; `District Offering of Full-Day Kindergarten` (mandatory / permissible / LEA option); `Year Full-Day Law Enacted`.
Derived variables	Years relative to mandatory attendance law adoption. Years relative to mandatory full-day offer law.

Attribute	Detail
Used in	Event-study around mandatory adoption ; pre-adoption trend check by partisan control; logit predictors of adoption.

Source 10: KlarnerPolitics — State Legislative Election Returns, 1967–2016

Attribute	Detail
Access	`196slers1967to2016_20180908.RData`; DOI: 10.7910/DVN/3WZFK9; loaded in `2_b_a_politics.R`.
Unit of observation	State × chamber × district × election year.
Geographic coverage	All 50 states (Nebraska unicameral treated separately).
Temporal coverage	Elections 1967–2016; the thesis uses the sub-period 1986–2016.
Variables extracted	State abbreviation (`sab`), state FIPS (`sfips`), chamber (senate / house), legislative district code, election year, winner party (Democrat / Republican / Other), vote count per candidate.
Derived variables	Using the "most-recent-winner-per-district" method: `Sen_dem`, `Sen_rep`, `Sen_other` (senate seat counts); `House_dem`, `House_rep`, `House_other` (house seat counts); `lean_senate` and `lean_house` (net Democratic partisan lean = (D seats − R seats) / total seats); `legis_control` (Democratic / Republican / Split).
For PA only	PA House and PA Senate electoral outcomes: winner party, Democratic vote share, signed Democratic margin, seat-flip indicator, turnout rate.
Used in	State-level TWFE correlates; logit of adoption; pre-adoption trend disaggregation by partisan control ; PA electoral outcomes.

Source 11: Governor Party Data — NCSL/NGA Historical Records

Attribute	Detail
Access	Compiled from the National Conference of State Legislatures (NCSL) and National Governors Association (NGA) historical records.
Unit of observation	State × gubernatorial term.
Geographic coverage	All 50 states.
Temporal coverage	Approximately 1979–2025.
Variables	`sab` (state abbreviation), `from` (start year of term), `to` (end year of term), `party` (Democrat / Republican / Independent).
Derived variables	`dem_gov`: binary indicator equal to 1 if the governor is a Democrat in a given year; `trifecta_lean`: composite full-government control indicator (-1 = Republican, 0 = Split, +1 = Democratic).
Used in	State-level TWFE correlates (outcome: *Dem governor (binary)*); logit of adoption (predictor: *Dem governor (pre-treatment)*).

Source 12: National Conference of State Legislatures — State Legislative Composition, 2009–2025

Attribute	Detail
Access	Files downloaded from NCSL; used for validation and to extend coverage beyond the Klarner dataset's 2016 endpoint.
Unit of observation	State × year.
Geographic coverage	All 50 states + D.C.
Temporal coverage	2009–2025 (one row per legislative session / year).
Variables	`State`; `Senate` (total seats); `Sen_dem`, `Sen_rep`, `Sen_other`; `House` (total seats); `House_dem`, `House_rep`, `House_other`; `Legis_control` (D / R / Split); `Governor` (D / R / I); `State_control` (trifecta indicator).
Used in	Validation of Klarner-derived seat counts for 2009; extension of the legislative composition panel to 2016–2024 for the TWFE.

Source 13: IPUMS CPS ASEC — State-Level Female labor Force Participation, Pre-2000

Attribute	Detail
Access	Microdata extract from IPUMS CPS (https://cps.ipums.org), cited in the thesis as "IPUMS CPS ASEC"; summary statistics computed and stored for the logit specification .
Unit of observation	State (cross-sectional average over pre-2000 years).
Geographic coverage	All 50 states + D.C.
Temporal coverage	Annual CPS ASEC surveys; averages computed over all years prior to 2000 for states that had not yet adopted mandatory KG by that point.
Variables extracted	State female labor force participation rate (share of adult women in the labor force or actively seeking work).
Derived variables	`female_lfp_std`: state pre-period female LFP rate, standardised to mean 0 and unit standard deviation, for comparability across states in the logit.
Used in	Logit predictors of mandatory FDK adoption (predictor: *Female LFP rate, std (pre-2000)*); interpretation of adoption endogeneity.

Public Policy Master's Thesis Series

This series presents the Master's theses in Public Policy and in European Affairs of the Sciences Po School of Public Affairs. It aims to promote high-standard research master's theses, relying on interdisciplinary analyses and leading to evidence-based policy recommendations.

Title

The Political Economy of Kindergarten Expansion

Name, Family name

Eve Margolis

Abstract

This thesis studies whether expanding full-day kindergarten (FDK) serves incumbents' electoral interests, combining a descriptive state-level analysis of adoption patterns across the United States with a causal district-level study of Pennsylvania. The causal design exploits staggered FDK adoption across school districts using a Callaway-Sant'Anna difference-in-differences framework. The descriptive analysis finds a small positive association between kindergarten enrollment intensity and Democratic legislative seat shares, with no detectable pre-trends in the political leaning of implementing states. The causal analysis finds that FDK adoption increases new voter registration among adults of plausible parenting age by approximately 0.33 percentage points per year, a 20–35 percent relative increase above baseline. This effect is present for both mothers and fathers, builds gradually over roughly a decade, and generates no detectable net partisan advantage, suggesting a civic mobilization mechanism rather than direct credit-claiming. A channel analysis finds positive but statistically imprecise labor force participation gains among single mothers of kindergarten-age children following adoption, suggestive of a time-constraint mechanism. On the electoral side, aggregate Democratic vote share and margins are negative and statistically significant at both chamber levels, while Democratic Senate incumbent retention shows a modest positive effect. The similarity between results regardless of the incumbency status suggests a non-partisan mobilization rather than a partisan dividend. Taken together, the findings suggest that FDK expansion is a civically meaningful policy whose political payoff accrues slowly and non-partisanly, and that targeted state subsidies toward low-wealth districts could help maximize welfare gains while improving the policy's long-run political sustainability.

Key words

Full-day kindergarten, political economy, voter registration, difference-in-differences, elections, labor force participation