THE EFFECTS OF THE FOUR-DAY SCHOOL WEEK ON TEACHER RECRUITMENT AND RETENTION

Andrew M. Camp[1 †], J. Cameron Anglum[2], Cory Koedel[3], Se Woong Lee[3] & Tuan D. Nguyen[3]

[1]Brown University
[2]Lehigh University
[3]University of Missouri

This version: February 2026
(Click here for most recent version)

ABSTRACT

School districts across the United States are increasingly adopting the four-day school week, often in the hope of addressing teacher staffing challenges. We examine the effects of the four-day week on teacher recruitment and retention in Missouri, where three in ten districts currently use it. After presenting qualitative data showing that school and district leaders, and teachers, believe the four-day week improves recruitment and retention, we estimate the effects of the four-day week on these outcomes using difference-in-differences models. We find no consistent evidence that the four-day school week improves teacher recruitment or retention. Our findings hold when considering potential saturation effects, indicating that even early adopters experienced no statistically significant benefit from adopting the four-day school week calendar.

Keywords: Teacher Turnover, Teacher Labor Markets, Four-Day School Week, Teacher Recruitment, Teacher Retention

Acknowledgements: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C240007 to American Institutes for Research. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

† Corresponding author: andrew_camp@brown.edu

1 INTRODUCTION

The four-day school week (4DSW) was first introduced in the United States in 1936 as a cost-saving measure. The policy failed to gain traction then, but it re-emerged with the Great Recession and by 2019, over 600 (primarily rural) U.S. school districts in 24 states were using the 4DSW (Thompson et al., 2021). Just four years later, in 2023, the number of districts using the 4DSW had increased to approximately 850 (Morton et al., 2024). Whereas the policy's pre-COVID expansion was motivated largely by anticipated cost savings—which often failed to materialize substantively (Kilburn et al., 2021; Morton, 2021; Thompson, 2021)—the motivation for more recent, post-COVID adoptions has increasingly centered around improving teacher recruitment and retention (Barnes & McKenzie, 2023; Camp, 2024).

Teacher recruitment and retention are particularly salient issues in rural school districts, which experience higher rates of teacher turnover and struggle to fill vacancies compared to districts in urban and suburban areas (Ingersoll & Tran, 2023; Rhinesmith et al., 2023). The 4DSW is often seen as a solution to address some of the specific problems rural districts face, such as longer teacher and student commute times. The 4DSW is also appealing to districts that are looking to be more competitive in the teacher labor market but face tight budget constraints, as it can be implemented without increasing expenditures and may be viewed as an attractive job characteristic by prospective new hires.

We contribute to the literature by studying the effects of the 4DSW on teacher recruitment and retention in Missouri. Districts in Missouri are among the most aggressive adopters of the 4DSW in the nation. The first Missouri school district switched to the 4DSW in the 2010-11 school year. By 2024-25, 178 districts (32 percent of all Missouri districts) had adopted the four-day calendar. In 2020, the Missouri State Board of Education acknowledged the shift toward the 4DSW in many districts and noted the potential for the policy to attract and retain teachers, especially in districts that struggle to offer competitive salaries. As documented by Turner et al. (2019), "Rural Missouri superintendents often struggle to compete in the teacher salary market; those that use the four-day school week model consider it an innovation that helps them attract and keep good teachers" (p.2).

While the 4DSW has expanded rapidly in Missouri, recent legislative proposals and public discussions have sought to curtail its expansion. Some have advocated for either the return to a traditional five-day schedule, or longer school years, due to concerns about the academic and workforce implications of reduced instructional days (Hanshaw, 2024). The costs and benefits of the 4DSW are actively debated in Missouri and elsewhere, but our understanding of its impacts on teacher recruitment and retention—the primary motivation for recent adoptions of the 4DSW—remains limited.

We begin our investigation by conducting semi-structured interviews with superintendents, principals, and teachers in Missouri school districts that have adopted the 4DSW. The interviews confirm that the primary rationale for the shorter week is to increase teacher recruitment and retention, particularly in the context of a perceived inability to compete on salary with other districts. Interviewees often rely on personal observations, such as increases in applicants for open positions and improved teacher satisfaction, as evidence of the 4DSW's impact.

Next, we estimate the effects of adopting the 4DSW on teacher recruitment and retention using difference-in-differences and event-study models. In contrast to the findings from our interviews, our quantitative analysis yields no evidence of practically meaningful 4DSW impacts. Our retention estimates are statistically insignificant and even if taken at face value, they are substantively small, implying the 4DSW reduces teacher exits by just 0.6 teachers per 100. We also find no evidence that the 4DSW boosts new-teacher recruitment, though there is a small shift in the composition of new hires toward those with prior experience in Missouri relative to new hires without Missouri experience, most of whom are novices. We also examine the potential for effect heterogeneity along several dimensions suggested by our interviews, where respondents sometimes emphasize staffing challenges in particular fields. However, we do not find evidence that the 4DSW affects staffing in hard-to-staff subjects, nor do we find evidence that adopting the 4DSW changes the prevalence of out-of-field teaching.

Perhaps the most interesting takeaway from our study is the disconnect between our qualitative and quantitative data—specifically, that the optimism about the 4DSW expressed in the interviews is not supported by our empirical analysis. This disconnect indicates a significant information problem and suggests better information could lead to better policy decisions. Addressing this information problem is especially important considering a separate but related body of research showing the 4DSW has mostly negative impacts on student outcomes.

2 BACKGROUND & PRIOR RESEARCH

The 4DSW is permissible due to policies in some states that allow school districts to meet instructional-time requirements in terms of instructional hours instead of school days. Not all states permit hours-based compliance, but in many states that do, the 4DSW has become more common. However, while the 4DSW is increasingly popular, it is also controversial. In Missouri specifically, a series of bills known as the "Parents' Bill of Rights," first introduced in 2022, attempted to mandate a return to the five-day school week for all Missouri districts (Ketz, 2023). Though the bills did not pass, a subsequent bill in 2024, Senate Bill 727 (SB727), did pass and requires districts in communities of at least 30,000 residents to secure voter approval via a ballot measure to continue using or adopting the 4DSW by Fall 2026. Additionally, SB727 authorized modest state funding increases—of one to two percent of prior state aid—for districts that maintain five-day schedules. Thus, while the 4DSW is widely used in Missouri, it is also the subject of active policy debate.[1]

The academic literature on the 4DSW has focused primarily on its effects along three dimensions: (1) district operating costs, (2) the teacher labor market, and (3) student outcomes. The first two represent intended policy outcomes—i.e., these are areas where districts are actively working to address a problem and the 4DSW is seen as a solution. The third has been treated mostly as indirect and incidental. While some districts suggest the 4DSW may improve student attendance, we are not aware of any 4DSW district where the first-order rationale for adopting the policy was to improve student academic outcomes directly.

Cost savings, as noted above, were the initial allure of the 4DSW and remain a rationale for some districts that adopt the 4DSW today (Anderson & Walker, 2015; Kilburn et al., 2021; Morton, 2021; Thompson, 2021; Thompson et al., 2021). However, research shows that the actual cost savings of the 4DSW are limited. For example, in Oklahoma, Morton (2021) estimates total district expenditures decline by about two percent of the average adopting district's budget, with savings concentrated in operations, transportation, and food services expenditures. Similarly, Thompson et al. (2021) use data from 24 states to study the financial implications of shifting to the 4DSW and find cost savings of one to two percent of total expenditures.

Improving teacher recruitment and retention is the more prevalent rationale for adopting the 4DSW, at least contemporarily (Anglum & Park, 2021; Kilburn et al., 2021; Thompson et al.,

[1] SB727 does not affect our analysis because it took effect after the last year of our evaluation period.

2021). Our analysis contributes to a small and mixed recent literature on how the 4DSW affects recruitment and retention. We are aware of just six state-level studies. In Arkansas, Camp (2024) finds a small positive effect of the 4DSW on teacher retention, and evaluations in Texas find somewhat larger positive effects (Lawson, 2025; Khalid, 2025). In Oregon, Ainsworth et al. (2024) find a negative retention effect. In Colorado, both Morton and Dewil (2024) as well as Bowser (2025) find null effects on teacher retention.[2]

Finally, we turn to student outcomes. Some advocates for the 4DSW suggest the calendar may improve student attendance by allowing families to schedule doctors' appointments or other family activities (e.g., helping with agricultural tasks) on the off day. However, most studies have estimated small and statistically insignificant impacts on average daily attendance (Anderson & Walker, 2015; Kilburn et al., 2021; Morton, 2023; Morton & Dewil, 2024), with one study finding that the 4DSW increased the share of students who were classified as chronically absent (Thompson et al., 2022). In terms of academic outcomes, most studies find the 4DSW reduces student achievement, though the negative achievement effects can only be statistically distinguished from zero about half the time (Kilburn et al., 2021; Morton, 2021; Morton et al., 2024; Thompson et al., 2022; Thompson & Ward, 2022).[3]

The effect sizes across studies range from -0.02 to -0.09 standard deviations of student test scores. While these may seem small at first glance, they reflect the effect of one year of exposure to the 4DSW and can cumulate into non-negligible learning loss over a student's schooling career.

[2] In a related study focusing on one suburban Colorado district, Nowak et al. (2023) find negative effects on teacher retention alongside negative impacts on student academic outcomes.

[3] There are two exceptions of which we are aware: Anderson and Walker (2015), who find positive effects of the 4-day week on proficiency rates in Colorado, and Morton (2023), who finds a mix of positive and negative achievement effects in Oklahoma high schools.

The negative achievement effects of the 4DSW suggest its implementation should be approached cautiously, especially if significant improvements along the intended dimensions are not clear.[4]

3 Motivation for District Adoptions of the Four-Day School Week

To provide context for our quantitative analysis, we conducted semi-structured interviews with 36 Missouri educators in 4DSW districts, which included 20 district superintendents, four school principals, and 12 teachers in 22 total districts. We focused primarily on superintendents, who play a key role in the decision to adopt and continue using the 4DSW. We recruited from rural districts across the state, including districts that have used the 4DSW for many years dating to the 2015-16 school year, and new adopters, as recently as the 2021-22 school year. All interviews were conducted via Zoom between October 2021 and April 2023, and all interviewees received a modest stipend. In light of the extant 4DSW literature, we followed a deductive coding strategy (Bingham & Witkowsky, 2022), seeking primarily to learn about district motivations for pursuing the 4DSW. These included school and district leaders' decision-making rationale regarding teacher recruitment and retention, financial considerations including teacher salaries and district cost savings, and the policy's appeal to teachers.

Interview respondents overwhelmingly cited improving teacher recruitment and retention as the primary objective of their 4DSW policies. District leaders often linked teacher recruitment and retention challenges to concerns about low teacher salaries and the inability to improve salaries. As one superintendent explained, "We can't compete with salary, so we had to put things in place like the four-day to encourage them to stay." Another commented on improving retention

[4] Thompson and Ward (2022) find that a key mediator of the negative effect of the 4-day week is time-in-school. They show that four-day districts with higher instructional hours experienced negligible achievement losses following policy adoption.

despite prevailingly low salaries: "They'll come here and work for two or three years and then they're going to go where they can make more money at the larger schools and I get that. But we're doing this to try to hold on to those folks a little bit longer." Along the same lines, one principal explained: "We're very close to districts that pay a lot more and so we're naturally at a disadvantage when trying to recruit any kind of staff member…The four-day week has become something that is pushed as a benefit." These views, which frame the 4DSW as an alternative to increasing salary, mirror similar superintendent interviews in Arkansas (Barnes & McKenzie, 2023).

The expansion of the 4DSW in Missouri has also brought on a flavor of "keeping up with the Joneses," where education leaders view it as a new dimension of competition in which they must compete. One respondent noted, "We're surrounded [by four-day districts] and if you're a teacher, you know, why would you pick us, right?...I know that we have lost some teachers to the four-day school weeks." And as another superintendent explained: "It becomes an expectation, or at least a norm, when you're trying to recruit teachers and you are competing for the same pool of people. The four-day school, it makes it tricky when you're not also a four-day school week and you're small, so small, relatively rural."

Although the quotes we have presented thus far represent the predominant view in our interviews, a small subset of district leaders noted challenges with using the 4DSW to improve recruitment and retention. For example, one superintendent hypothesized diminishing returns to 4DSWs: "I think if everybody does [the 4DSW], then what's going to happen is our recruitment advantage is going to disappear. I do believe that." Even more directly, another superintendent lamented using 4DSWs in lieu of other reforms she thought might be more sustainable: "If it comes to pass that the 4DSW becomes so ubiquitous that everybody has [it], then where are we going to be in rural America trying to recruit teachers? Or, we're going to have to find exactly what we should be doing now, which is building on a climate and culture that is more inviting for whatever reason. Some people prefer a small-town atmosphere over an urban atmosphere. That should be our rallying cry. Not, 'we will make you work for less, fewer days in the week.'"

These counterexamples notwithstanding, our interviews broadly confirm that most school and district leaders who participated in our study feel that the 4DSW has positive impacts on teacher recruitment and retention. Citing surveys of his teachers, one superintendent indicated the 4DSW represents "a life and career changing thing" for some. Another stated "We had more applications for our positions than we've ever had, since I've been here in the last four years. Our turnover was lower than it has been the last 10 years." A third commented on reversing teacher turnover and offsetting difficulties caused by pay discrepancies: "We had several teachers that left for higher pay that came back, and they were teachers we wanted to have here."

We also interviewed teachers directly. The teachers we interviewed generally reinforced district and school leaders' positive views of the 4DSW. Recalling how she navigated a district move, an elementary science teacher explained, "A four-day school week just kind of sweetened the deal for me. I was like, I remember the difference between a five-day and four-day for me in my mind, that sounds a little bit better to have that extra day." Similarly, a high school English teacher summarized, "The majority of the staff members where I teach are very pleased to have the four-day week…I know some people, that gives them extra [time] on some things for school work, or just to have that work-life balance." Even teachers otherwise committed to remaining in their districts noted marginal benefits of the policy. For instance, a high school math teacher indicated, "I would still stay there even if we [had] remained on a five-day. I will say, though, it has in a lot of ways made my made my job a little bit more enjoyable… I personally think the four-day week has encouraged a lot of our staff to stay, enjoy the extra day off."

In sum, district and school leaders, as well as teachers themselves, expressed positive views about the potential for the 4DSW to improve teacher recruitment and retention during our interviews, often drawing on their own personal experiences. Though there were a handful of dissenters, they were small in number and mostly questioned the policy's sustainability. While we do not claim that interviewed districts are representative of the state as a whole, their motivations for adopting a 4DSW calendar align with what state policymakers have called the "heart of the issue"—teacher retention and recruitment (Hanshaw, 2024).

4 Data

Our quantitative analysis is based on individual-level administrative panel data provided by the Missouri Department of Elementary and Secondary Education (DESE). Our data panel begins with the 2008-09 school year—two years before the first Missouri district adopted the 4DSW—and runs through the 2023-24 school year. These data include information about each teacher's school and district assignment, workload, experience, education level, certification area, teaching assignment(s), and basic demographic information (race/ethnicity and gender). We identify teachers based on position codes contained in DESE's data and restrict our sample to full-time employees, which we define as a full-time equivalency of 0.75 or more.[5]

[5] District FTEs were calculated across both teaching and non-teaching positions. Individuals that held both a teaching and non-teaching position were considered teachers. For the very small number of teachers assigned to multiple districts, we randomly selected one assignment per individual. Our results are qualitatively similar when including individuals with a full-time equivalency of less than 0.75.

At the district level, we merge in annual information on total enrollment, the demographics of the student population, urbanicity, charter status, and poverty proxied by the share of students enrolled in the free- and reduced-priced lunch program, all obtained through the National Center for Education Statistic's Common Core of Data. Our data on school districts' adoptions of the 4DSW are from DESE.[6] These data run through the 2024-25 school year, one year past the end of our administrative data panel. Figure 1 illustrates the expansion of the 4DSW in Missouri. By 2024-25, 178 of Missouri's 559 districts had adopted the policy.

Table 1 provides summary statistics for our sample. The values are district-level averages weighted by teacher full-time equivalents (FTEs). Panel A compares districts that ever adopted the 4DSW to districts that never adopted the 4DSW, using data from 2023-24. The comparison reveals, unsurprisingly, that adopters of the 4DSW are disproportionately rural, which comes with a host of correlated characteristics such as lower enrollment, fewer teachers, larger shares of White students, and lower average teacher pay, relative to districts that never adopt the 4DSW.

In panel B, we restrict the sample to districts that ever adopted the 4DSW, then compare early and late adopters, defined by whether districts adopted the 4DSW by 2020-21. We select this school year to separate early and late adopters because of the large increase in adoption following the 2020-21 school year, as shown in Figure 1. However, we note that this distinction is made only to highlight the relative similarities among adopted and not-yet-adopted districts; the comparison is qualitatively similar if we choose a different year to differentiate early vs. late adopters. Relative

[6] We cross-checked the DESE list with a separate list maintained by Dr. Jon Turner at Missouri State University and resolved a handful of discrepancies. We also found that two districts adopted the 4-day week but reverted back to a traditional five-day week by the end our data panel. We consider these districts as "always treated" in our analysis, but our findings are also insensitive to dropping them.

to the differences in panel A, the differences in panel B are much smaller, and most are not statistically significant.[7]

Our identification strategy relies on a difference-in-differences approach, which we describe in detail below, and it does not require level equivalence between the treatment and control groups at baseline for identification. Still, the large differences in panel A of Table 1 raise concerns about using never-treated districts as controls for districts that adopt the 4DSW. Thus, we use districts that have not yet adopted the 4DSW as the comparison group for adopting districts in any given year—i.e., we compare treated districts to not-yet-treated districts, as is commonly advocated in the recent difference-in-differences literature in similar circumstances (Callaway & Sant'Anna, 2021; Roth et al., 2023).[8]

5 METHODS

The difference-in-differences literature has expanded rapidly in recent years. To illustrate our identification strategy, it is instructive to begin with a conventional model. Focusing first on our analysis of teacher retention, consider the following textbook two-way fixed effects specification:

$$Y_{ijt} = \beta_0 + 4DSW_{jt+1}\beta_1 + \Phi_j + \zeta_t + \eta_{ijt} \qquad (1)$$

In equation (1), Y_{ijt} is an indicator equal to 1 if teacher i, working in district j in year t, exits the district after year t (i.e., is not present in the district in year $t+1$). We also separate exits by the type of exits, where there are two groups: teachers who move to other Missouri districts and

[7] The differences in panel B would be even smaller if we omitted one large, suburban, late-adopting district (Independence School District). This district accounts for approximately 21% of the teacher-weighted averages among late adopters (e.g., the entire "suburban" share among late adopters is from this single district). In results available upon request, we confirmed our results presented below are similar if we exclude this district from our analysis entirely.

[8] Below and in Appendix C we provide additional evidence to support this decision.

teachers who leave the Missouri teaching workforce (i.e., transfer to a district in another state or to a private school, or exit the teaching profession). The variable $4DSW_{jt+1}$ is an indicator equal to one if district j used the 4DSW in year $t+1$ and zero otherwise. District and year fixed effects are denoted by Φ_j and ζ_t, respectively, and η_{ijt} is an idiosyncratic error. With the inclusion of district and year fixed effects, β_1 is a difference-in-differences estimate of the average effect of the 4DSW on teacher retention. We do not include control variables because they are not necessary for identification under the difference-in-differences assumptions and their use requires imposing assumptions about the effect of treatment on the evolution of these variables (Sant'Anna & Zhao, 2020).[9] Standard errors are clustered at the district level in all models.

Equation (1) identifies the causal effect of the 4DSW under the assumption that the treatment and comparison groups would have followed similar outcome trends in the absence of treatment. This is commonly referred to as the parallel trends assumption. The recent difference-in-differences literature has raised concerns about two assumptions subsumed by parallel trends. Both assumptions are about treatment-effect heterogeneity and are of concern because in addition to using untreated districts as control observations, equation (1) also uses earlier, already-treated districts as control observations for more-recently treated districts. If treatment effects are homogenous then already-treated observations are appropriate controls and can satisfy the parallel trends assumption. However, if there is treatment-effect heterogeneity, it could introduce bias.

The first type of treatment effect heterogeneity that could cause bias is with respect to treatment duration, which we refer to as duration heterogeneity. As an example in our context, teachers may become increasingly convinced of the permanency of the 4DSW as the policy

[9] For completeness, in Appendix A we re-estimate our models after including teacher and district-level characteristics as control variables. Our findings are substantively similar (see Appendix Table A.2).

endures in district j, leading to a stronger retentive effect the longer the policy is in place. If this type of heterogeneity exists, it would be problematic because equation (1) will compare new 4DSW adopters to older adopters, whose retention trends would be improving due to duration heterogeneity, violating parallel trends.

The second type of treatment effect heterogeneity is with respect to when districts adopt the policy, independent of duration heterogeneity. We refer to this type of heterogeneity as group heterogeneity. Group heterogeneity may arise in our context if, for example, the effect of the 4-day week on retention is larger for early adopters when the policy was rare in Missouri, making it a more unique amenity. Groups of later adopters would face a more crowded market, potentially dulling the effect of the 4DSW. This type of heterogeneity would also be problematic in equation (1) due to the model's use of earlier-treated units as control observations, following the same logic as above.

In the presence of either type of heterogeneity, equation (1) can result in misleading inference through the comparison of treated units to each other when the treated units differ by treatment duration and group. Roth et al. (2023) describe these as "forbidden comparisons." The literature has proposed a number of solutions to these issues in difference-in-differences estimation (e.g., Borusyak et al., 2024; Callaway & Sant'Anna, 2021; De Chaisemartin & D'Haultfœuille, 2020; Gardner et al., 2024; Goodman-Bacon, 2021; Sun & Abraham, 2021; Wooldridge, 2021). Our preferred models use the approach of Gardner et al. (2024).

Gardner et al. (2024) split the traditional difference-in-differences specification into two stages and in doing so, mechanically block forbidden comparisons. To implement their approach, in the first stage we estimate the following regression using all district-year observations when the 4DSW is off—i.e., not-yet-treated districts, where $4DSW_{jt+1} = 0$:

$$Y_{ijt} = \alpha_0 + \theta_j + \delta_t + \nu_{ijt} \quad (2)$$

Equation (2) is similar to equation (1) but excludes the treatment variable ($4DSW_{jt}$) and is estimated on the restricted sample. θ_j and δ_t are district and year fixed effects, as in equation (1). We use the estimates from equation (2) to construct the following residualized values of the outcome variable in the full sample in all years, as follows:

$$Y^*_{ijt} = Y_{ijt} - \left(\hat{\alpha}_0 + \hat{\theta}_j + \hat{\delta}_t\right) \quad (3)$$

We then use these residuals as outcomes in the following second-stage regression, estimated on the full sample:

$$Y^*_{ijt} = \gamma_0 + 4\text{DSW}_{jt}\gamma_1 + \varepsilon_{ijt} \quad (4)$$

The key differentiating feature of the Gardner et al. approach—in fact, the only thing that distinguishes it substantively from the standard difference-in-differences model—is that the district and year fixed effects, θ_j and δ_t in equation (2), are estimated using only untreated observations in the first stage (or, in our application, not-yet-treated observations). This mitigates the potentially contaminating effects of duration and group heterogeneity by preventing already-treated districts from contributing to the control function. Garnder et al. (2024) show this approach produces results similar to other common approaches that address these same issues, with the benefits that the two-stage approach is less computationally burdensome and (generally) more precise. In Appendix A, we confirm the similarity of our estimates to analogous estimates from other approaches commonly used in the literature (Borusyak et al., 2024; Callaway & Sant'Anna, 2021; Wooldridge, 2021).

Like any difference-in-differences estimate, γ_1 is the average treatment effect over groups and durations observed in the actual data. For instance, if duration effects are important, and the sample is tilted toward long-term adopters, this will influence the average treatment effect. This is appropriate as it would reflect the treatment experience in the sample. One way to isolate duration effect heterogeneity is to estimate the model as an event study. Extending the model in this way is straightforward and simply requires replacing the treatment indicator in equation (4) with a vector of leads and lags for policy adoption (Gardner et al., 2024). The event-study framework also allows us to assess the plausibility of the parallel trends assumption by examining the coefficients on the pre-treatment indicators. Though the parallel trends assumption cannot be tested directly because the post-treatment counterfactual trends are unobserved, it is common to provide evidence on its plausibility by testing for parallel trends leading up to treatment (i.e., parallel pre-trends). A statistically significant pre-trend would raise the concern that the identifying assumption is unlikely to be met.

Our approach to studying teacher recruitment is methodologically similar, but with the complication that new teachers cannot be assigned to specific districts (and thus a treatment condition) before they are hired. We address this limitation by aggregating our data to the district level and testing whether districts that adopt the 4DSW are able to hire more new teachers. Consider the following district-aggregated analog to equation (1):

$$Y_{jt} = \lambda_0 + 4DSW_{jt}\lambda_1 + \psi_j + \pi_t + u_{jt} \qquad (5)$$

In equation (5), Y_{jt} is the new-teacher share of total teacher FTE in district j in year t, and ψ_j and π_t are district and year fixed effects, respectively. We weight equation (5) by each district's teacher FTE in year jt to ensure our findings are comparable to our teacher-level retention findings (without the weights, the district-level model would implicitly give more weight to smaller districts), and estimate the model using Gardner et al.'s two-stage approach as illustrated above.

We construct versions of Y_{jt} based on all new teachers to the district and also separate new teachers into groups who (a) move from other Missouri districts and (b) are new to teaching in Missouri.

The district-level aggregation of the data needed to facilitate our analysis of teacher recruitment is not very costly. This is because the policy variation is at the district level. With district-level clustering, the teacher-level retention analysis is also effectively conducted at the district level, and in fact the results from our retention analysis are very similar if we aggregate the data to the district level (see Appendix Table A.1).

We conclude with an acknowledgment that our analysis of new teachers is an imperfect analysis of teacher recruitment *ex ante* because the number of new teachers will depend on other aspects of the teacher labor market. Most notably, it will depend on the number of vacancies, which depends in part on retention, and retention may depend on the 4DSW. However, as we discuss below, because we do not find evidence of retention effects of the 4DSW, our findings are informative about the recruitment channel *ex post*.

6 RESULTS

We begin in Table 2 with results from our preferred retention models shown in equation (4). Row (1) shows results from the model where the outcome is any turnover. Rows (2) and (3) separate exits by type: row (2) shows results for moves to a different district and row (3) shows results for exits from the workforce.[10] Reading across the columns, we estimate the models on different teacher samples. In Column (1) we use the full teacher sample and in column (2) we use high-need

[10] We cannot disentangle teachers who exit the teaching profession from cross-state movers in the administrative data, which is a typical problem with state data systems. However, in studies that merge multiple state data systems and are able to disentangle these different types of moves, the results suggest the vast majority of exits from the state data system are likely exits from the teaching profession altogether (Goldhaber et al., 2015; Podgursky et al., 2016).

teachers, who we define as teachers in STEM fields, STEM-adjacent fields (e.g., career and technical education and applied-science fields), and special education. In columns (3) and (4) we separate out teachers in STEM and special education, respectively.

Focusing first on our primary estimate in row (1) and column (1), we find no evidence that adopting the 4DSW reduces teacher exits. Our statistically insignificant point estimate, if taken at face value, is also very small—it implies that the 4DSW results in just 0.6 fewer teacher exits per 100 teacher FTEs. Furthermore, columns (2) to (4) give no indication the effect of the 4DSW is stronger among teachers in high-need areas. None of the coefficients in these columns is statistically significant, and the individual estimates are inconsistent in sign. One might be tempted to draw inference from the suggestive coefficient for STEM teachers in row (1). However, even this coefficient is not statistically significant and taken at face value, its substantive implications are small (i.e., it suggests the 4DSW leads to 1.3 fewer STEM teacher exits per 100 teacher FTEs). We also echo the caution given by von Hippel and Schuetze (2025) about drawing inference from this type of "no, but" heterogeneity in instances where there is not a detectable main effect in the full sample.

Figure 2 provides event-study analogs to our full-sample estimates in column (1) of Table 2. The event studies show strong evidence of parallel pre-trends for all outcomes, with small confidence intervals and point estimates near zero. There is not even a suggestive trend in the post-treatment event study parameters to suggest a 4DSW effect.

We test the robustness of our findings to several alternative specifications. First, in Appendix Table A.1, we confirm our results are nearly identical if we estimate our retention models using district-aggregated data, which matches the analytic structure of our recruitment analysis that follows. Second, in Appendix Table A.2, we add covariates within the Gardner et al.

framework. Our coefficients change very little, though the estimate for STEM teachers changes from insignificant to marginally significant at the 10 percent level. Third, in Appendix Tables A.3, A.4, and A.5, we show that our findings are substantively similar if we use the alternative difference-in-differences estimators proposed by Borusyak et al. (2024), Wooldridge (2021), and Callaway and Sant'Anna (2021).

Next, we turn to our analysis of teacher recruitment. Table 3 shows estimates using Gardner et al.'s two-stage method. The structure of the table follows from Table 2. The first row shows estimates for all new teachers regardless of where they came from, and rows (2) and (3) divide new teachers into two groups based on whether they came from a different Missouri district or are new to Missouri. Column (1) gives our estimates for new teachers in all subjects, and columns (2) to (4) show results separately for new teachers whose initial assignments are in high-need fields.

As noted above, a challenge with interpreting our findings for teacher recruitment is that the recruitment of teachers depends on vacancies. So, for instance, if the 4DSW had large and positive retention effects, we might expect fewer new teachers to appear in 4DSW districts, but this would not necessarily mean the policy is ineffective. However, given our very small and statistically insignificant retention estimates, this type of confounding is largely ruled out. Conditional on our null findings for retention in Table 2, our results in Table 3 speak to whether the 4DSW improves teacher supply by sparking new interest from outside the district.

Our results in row (1) of Table 3 give no indication of a recruitment benefit of the 4DSW overall, nor for any of the subgroups that we examine (i.e., teachers in high-needs settings, STEM teachers, and SPED teachers). Complementary event-study graphs in Figure 3 point to the same conclusion. This result is perhaps as expected if adopters of the 4DSW are fully staffed. However, if part of the impetus for adopting the 4DSW is to address staffing shortages, and the 4DSW is effective in this regard, we would have expected an increase in the new-teacher FTE share as more teachers are drawn to the district with the adoption of the policy.

In rows (2) and (3) of Table 3, we do find evidence of a small effect on the composition of new hires: districts that adopt the 4DSW attract relatively more teachers from other Missouri districts and fewer teachers new to Missouri. These effects are offsetting and the shift is on the order of 1 teacher FTE per 100. These findings are substantively similar to what Lawson (2025) finds in Texas, although the composition shift in Missouri is much smaller.

7 Extensions: Labor Market Stress Indicators, Spillovers, And Saturation

We extend our analysis in three ways. First, we apply the same aggregated model we use for our analysis of teacher recruitment to examine "stress indicators" of the teacher labor market. Our stress indicators capture the FTE shares of teachers with STEM, STEM-adjacent, and special education teaching assignments who are not certified to teach in these fields.[11] If the 4DSW addresses staffing challenges in a subtle way that is missed by measures of teacher FTE, its effect could show up in the form of a reduced reliance on out-of-field teachers in high-need areas. The results are reported in Table 4 and give no indication that the 4DSW reduces out-of-field teaching.

Second, we test whether spillover effects across districts could bias our findings. Our approach follows Butts (2024), who shows that in the presence of spillover effects, the standard difference-in-differences parameter (using any credible methodology) is an estimate of the average treatment effect minus the average spillover effect on untreated districts. He shows that with modest modifications to the Gardner et al. (2024) approach, and under the assumption that

[11] We identify teachers as out-of-field if they are not certified at all or if they are certified but the certification subject area does not match their teaching assignment. The use of completely uncertified teachers in Missouri is very uncommon and most out-of-field teachers fall into the latter category.

spillovers are localized to some degree, bias from spillover effects can be removed from the difference-in-differences parameter.

To illustrate the approach, we momentarily assume there is some distance, d, over which the effect of the 4DSW in one district spills over to another. Beyond d, there is no spillover effect. If we know d, we can re-estimate equation (2), but in addition to excluding district-years when the 4DSW is in effect, we can also exclude district-years where another district within distance d has adopted the 4DSW (i.e., any district-year subject to spillovers). The rationale for this exclusion in the first stage is the same as the rationale for excluding treated districts in our primary analysis, in this case applied to districts that are affected by the 4DSW through spillovers.

We can then use the residual values (from equation 3) to estimate the following second-stage regression, which is analogous to equation (4):

$$Y_{ijt}^{*} = \omega_0 + 4\text{DSW}_{jt}\omega_1 + Spillover_{jt}^{d}\omega_2 + \epsilon_{ijt} \qquad (6)$$

Equation (6) differs from equation (4) only in that it includes the term $Spillover_{jt}^{d}$, which is an indicator variable equal to one if district j is not directly treated but subject to spillovers—i.e., it has not adopted the 4DSW itself, but is exposed to at least one other district within d kilometers that has adopted the 4-day week. The $Spillover_{jt}^{d}$ variable effectively separates the control group into two smaller groups: one that is affected by spillovers, and one that is not, with the latter omitted. This forces the difference-in-differences parameter, ω_1, to identify the effect of the 4DSW by comparing treated observations to control observations that are not contaminated by spillovers. If there is bias in our primary estimates from equation (4) due to spillover effects, then the estimates from equation (6) should be different because they will not include the bias.

To implement this test for bias, we must specify d. While a rich literature shows that the teacher labor market is highly localized (Boyd et al., 2005; Edwards et al., 2023; Engel & Cannata, 2015; Reininger, 2012), and in results suppressed for brevity we confirm this is true in Missouri, there is no analytic way to precisely define d.[12] Therefore, we show a range of estimates from equation (6) over integer values of d from [0, 150] km (corresponding to a range of 0-93 miles). When $d = 0$, equation (6) replicates equation (4) and as d grows, it removes bias from potential spillovers under different assumptions about their geographic reach.

Figure 4 shows the results from this exercise for our retention models, focusing on the main estimates in column (1) of Table 2. Each graph in the figure plots 151 different estimates of ω_1, corresponding to the estimation of the spillover model over integer values of $d \in [0,150]$. The primary takeaway from the figure is that there is no indication that spillover effects are causing bias in our primary models.

Figure 5 shows analogous results from our recruitment analysis, corresponding to the estimates in column (1) of Table 3. Visually, the offsetting estimates for new teachers from other districts, and new teachers from outside of Missouri, converge toward zero at high values of d, but the error bands include the original estimates at virtually all values of d in all models. We again conclude there is no evidence of meaningful bias from spillover effects.

Next, we consider the related issue of potential effect heterogeneity caused by market saturation—i.e., spillover effects on treated districts from other treated districts. The results we have presented thus far are average treatment effects for adopting districts across all adopting

[12] For example, among observed teacher moves across district lines, half occur within 41 km (26 miles).

cohorts and years. However, if the effect of the 4DSW depends on how common it is, we might expect different effects depending on the prevalence of the policy.

To evaluate the potential for saturation effects, we begin by constructing a measure of 4DSW saturation for each adopting district in its local area at the time of adoption. We use commuting zones (CZs) to define the local area (Fowler, 2024). CZs are groups of contiguous counties defined based on commuting flows from home to work as reported in the American Community Survey and have been used in prior research to approximate local labor markets, including teacher labor markets (Autor & Dorn, 2013; Edwards et al., 2025). For each treated district, we measure local-area 4DSW saturation at the time of adoption by the share of total teacher FTEs in the CZ in 4DSW districts, excluding FTEs in the adopting district itself.

We incorporate market saturation into our models by augmenting equation (4). Specifically, we interact the treatment indicator with our continuous measure of the prevalence of the 4DSW in commuting zone z at the time district j adopts the 4DSW, t^*, as follows:

$$Y^*_{ijzt} = \rho_0 + 4\text{DSW}_{\text{jt}}\rho_1 + \left(4\text{DSW}_{jt} \times \text{P}_{zjt^*}\right)\rho_2 + \epsilon_{ijzt} \tag{7}$$

Equation (7) leverages variation in the prevalence of the 4DSW across CZs and over time to identify the effect of adopting the 4DSW in more and less saturated markets.[13] The parameter ρ_1 represents the effect of 4DSW calendar adoption net of saturation (i.e., when the prevalence of the 4DSW in the CZ is zero), and ρ_2 represents the effect of saturation. If, for instance, we found a negative coefficient on ρ_1 and a positive coefficient on ρ_2 in our model of teacher turnover, this would indicate short-term benefits of the 4DSW that attenuate as more nearby districts adopt the policy.

---

[13] See Appendix B for more details about the temporal and spatial variation.

Tables 5 and 6 present our findings from equation (7) for teacher turnover and recruitment, respectively. There is no indication that the effect of the 4DSW is moderated by saturation. This is evident in our findings in two ways. First, the coefficients on the primary 4DSW treatment variable—i.e., ρ_1—are very similar to what we report in Tables 2 and 3 when we do not account for saturation. Second, the coefficients on the interaction with saturation are consistently small and far from statistical significance. While these coefficients may initially seem large, this is only because they measure the effect of hypothetically going from 0-percent to 100-percent saturation (i.e., from none to all teacher FTEs in the CZ working in 4DSW districts). Noting that the standard deviation of the saturation variable across years and CZs is just 0.10, the coefficients indicate the scope for saturation to influence the treatment effect is very small.

In Appendix B, we report event study analogs of equation (7) for both teacher turnover and recruitment, which corroborate this interpretation.

8 OTHER CONSIDERATIONS

In this section we briefly address several other considerations related to our analysis and findings. First is our decision to use not-yet-treated districts as the control group. This is an intuitive strategy commonly used in modern difference-in-differences applications and motivated by the large level differences between ever- and never-treated districts in Table 1. In short, never-treated districts are not a natural comparison group for districts that ultimately adopt the 4DSW in Missouri. In Appendix C, we explore this issue further by considering two alternative control groups: (1) a group that consists of not-yet-treated districts plus rural never-treated districts, and (2) a group that includes all untreated districts—i.e., all never and not-yet-treated districts.

We plot event-study trends for six different district characteristics—student enrollment, share of non-white students, share of students enrolled for free and reduced-price lunch, average teacher experience, share of teachers with a master's degree or more, and residualized teacher salary (conditional on a cubic of experience and teacher education)—using our preferred control group of not-yet-treated districts and the two alternatives. Using our preferred control group, and the control group that also includes rural never-treated districts, there is no evidence of pre-trends for any of the six district characteristics. However, when we add all never-treated districts to the control group, pre-trends emerge for two of the six characteristics—student enrollment and the share of students enrolled for free and reduced-price lunch. Combined with the clear non-comparability between ever- and never-treated districts documented in Table 1, these trends reinforce our decision to exclude non-rural, never-treated districts from our main models. The pre-trends for the other two control groups give no cause for concern. Correspondingly, in the appendix we show that if we expand our control group to include rural never-treated districts, our results are substantively similar to what we report above.

A related issue addressed in Appendix C is whether other aspects of school districts change after the adoption of the 4DSW. Any such changes post-policy could be informative about effect moderators. A notable test is for teacher salaries because in Oregon, Ainsworth et al. (2024) find that adoption of the 4DSW is accompanied by slower growth in teacher salaries, relative non-4DSW districts. This may offset the benefit of the policy and help to explain their finding that it is associated with lower teacher retention. However, unlike in Oregon, Appendix Figure C.6 shows no evidence that adopting the 4DSW in Missouri is accompanied by changes to districts' salary profiles. Moreover, the clean pre-treatment trend further suggests the not-yet-treated group of districts is serving as an appropriate control. We conclude that our null findings for the labor market effects of the 4DSW are not influenced by concurrent changes in teacher salaries in adopting districts.

We similarly examine how average teacher experience and education evolve in districts that adopt the 4DSW (in Appendix Figures C.4 and C.5). We do not find significant experience or education effects, though there is the suggestion of a short-term boost to both. This could reflect the composition shift in new hires we identify in Table 3—i.e., from teachers with no prior Missouri experience to teachers from other Missouri districts—but we do not interpret these results strongly because they are not significant and in most years, they are quite close to zero.

Next, we reconsider the contrast between the positive responses during our interviews with superintendents, principals, and teachers, and our largely insignificant quantitative results. While Occam's razor suggests the most likely explanation is that the perceptions of the 4DWS do not match the empirical reality, another possibility is that the districts we interviewed just happened to have more success with the 4DSW policy than the average district. If so, interviewees in the subset of districts we interviewed could have correct perceptions in their districts, even if the larger effects of the policy are null.

To test for this possibility, in Appendix D we re-estimate our main models using only the 22 districts represented in the interview sample as treated. That is, we drop all non-interviewed adopters of the 4DSW from the treatment sample. We still include all not-yet-treated districts in the control pool to maintain a common baseline for comparison. A limitation of this analysis is that with the small treatment sample our estimates are less precise, but our findings are very similar. There is no indication that the interview districts have had better success with the 4DSW than the average adopter in Missouri.

Finally, there is the issue of the COVID pandemic. Our primary analysis takes no unique consideration of the pandemic period, instead relying on year fixed effects to absorb its general impact. But it is possible that the disruption of the pandemic could have dulled the retention and recruitment effects of the 4DSW. To test for this possibility, we replicate our main models but restrict the sample to pre-COVID years (specifically, we only include retention and new hires through the 2019-20 school year). This analysis also complements our analysis of saturation effects above, in that these early adopters were in a less-saturated environment. With the same caveat above about the loss of precision due to a reduced sample size, our results in Appendix E show no evidence that the 4DSW had larger effects prior to the pandemic when the policy was less common in Missouri.

9 CONCLUSION

The four-day school week (4DSW) continues to gain traction across the U.S., especially in rural school districts. While the prevailing rationale is that the 4DSW is a tool to improve teacher recruitment and retention, the research literature on how the 4DSW affects these outcomes is thin. We contribute to this emerging literature by studying the 4DSW in Missouri, a state where the policy is widely implemented. We find no evidence that the 4DSW improves teacher recruitment or retention overall, though we document a small shift in the composition of new hires in 4DSW districts toward teachers with prior experience in Missouri.

Although our mostly null findings may initially seem surprising, the existing literature provides limited support for the claim that the 4DSW improves teacher recruitment and retention. Camp (2024) finds a small positive effect on retention in Arkansas, but both Morton and Dewil (2024) and Bowser (2025) find null effects on recruitment and retention in Colorado, while Ainsworth et al. (2024) and Nowak et al. (2023) find negative retention effects in Oregon and Denver, respectively. To the best of our knowledge, the only state where meaningful positive effects of the 4DSW have been detected is Texas (in two studies: Khalid, 2025, and Lawson, 2025).

The contrast between our qualitative interviews and quantitative findings in Missouri raises an important question: how did the positive narrative around the 4DSW emerge among educators? While we do not claim to have definitive answers, we identify three factors that we believe have contributed to their positive views of the 4DSW. First, our findings do not necessarily mean that teachers do not value the 4DSW—only that they do not value it enough to significantly influence their employment decisions. Indeed, one teacher we interviewed suggested as much, indicating the 4DSW did not change the employment decision, but "made [the] job a little bit more enjoyable." Second, our interviewees lack a clear counterfactual for their experiences. It is possible that labor market conditions improved in districts adopting the 4DSW, but also improved in non-adopting districts, which would reconcile the null effects in our models with the locally perceived gains. Third, confirmation bias may play a role. Education leaders in 4DSW districts may be especially attuned to anecdotal successes, while unintentionally overlooking instances where the policy had no effect or even a negative effect on teacher recruitment or retention. Regardless of the source of the discrepancy, our findings highlight a substantial gap between perception and reality. This is especially concerning given that most prior research shows the 4DSW harms students' academic outcomes (Kilburn et al., 2021; Morton, 2021; Morton et al., 2024; Thompson et al., 2022; Thompson & Ward, 2022).

While the 4DSW continues to expand, resistance is growing, and our analysis contributes to active policy debates on this issue in Missouri and elsewhere. The resistance in Missouri is exemplified by Senate Bill 727, which requires large communities to obtain voter approval before adopting the 4DSW and provides additional funding to districts that retain a five-day schedule. A more proactive policy approach would involve addressing the underlying issue that has driven the expansion of the 4DSW of late: concerns about teacher recruitment and retention (primarily) in rural districts. Providing resources and supports for districts to address their staffing challenges, ideally in ways that do not harm student achievement, could facilitate a natural transition away from the 4DSW.

10 REFERENCES

Ainsworth, A. J., Penner, E. K., & Liu, Y. (2024). Less is more? The causal effect of four-day school weeks on employee turnover. (Working Paper 24-1035; EdWorkingPapers). https://doi.org/10.26300/QGTJ-MJ51

Anderson, D. M., & Walker, M. B. (2015). Does shortening the school week impact student performance? Evidence from the four-day school week. *Education Finance and Policy*, *10*(3), 314–349. https://doi.org/10.1162/EDFP_a_00165

Anglum, J. C., & Park, A. (2021). Keeping up with the joneses: District adoption of the 4-day school week in rural Missouri. *AERA Open*, *7*, 233285842110028. https://doi.org/10.1177/23328584211002842

Barnes, K., & McKenzie, S. (2023). We wanted to do something innovative (Working Paper 20–1; Arkansas Education Reports). Office for Education Policy. https://oep.uark.edu/we-wanted-to-do-something-innovative/

Bingham A. J., Witkowsky P. (2022). Deductive and inductive approaches to qualitative data analysis. In Vanover C., Mihas P., Saldaña J. (Eds), Analyzing and interpreting qualitative data: After the interview. Sage Publications.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, *91*(6), 3253–3285. https://doi.org/10.1093/restud/rdae007

Bowser, K. M. (2025). The four-day gamble: The quasi-experimental effects of four-day school week adoption on teacher, principal, and paraprofessional staff turnover and district financial outcomes (Working Paper 25-1208; EdWorkingPapers). https://edworkingpapers.com/ai25-1208

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management*, *24*(1), 113–132. https://doi.org/10.1002/pam.20072

Butts, K. (2024). Difference-in-differences estimation with spatial spillovers (Version 3). arXiv. https://doi.org/10.48550/ARXIV.2105.03737

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Camp, A. (2024). Teacher retention and quality in the four-day school week (Working Paper 24–971; EdWorkingPapers). https://edworkingpapers.com/ai24-971

De Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/10.1257/aer.20181169

Edwards, D. S., Kraft, M. A., Christian, A., & Candelaria, C. A. (2023). Teacher shortages: A unifying framework for understanding and predicting vacancies (Working Paper 22–684; EdWorkingPapers). https://www.edworkingpapers.com/ai22-684

Engel, M., & Cannata, M. (2015). Localism and teacher labor markets: How geography and decision making may contribute to inequality. *Peabody Journal of Education*, *90*(1), 84–92.

Fowler, C.S. (2024). New Commuting Zone delineation for the U.S. based on 2020 data. *Scientific Data* 11, 975.

Gardner, J., Thakral, N., Tô, L. T., & Yap, L. (2024). Two-Stage differences in differences. Working Paper.

Goldhaber, D., Grout, C., Holden, K. L., & Brown, N. (2015). Crossing the border? Exploring the cross-state mobility of the teacher workforce. *Educational Researcher*, *44*(8), 421–431. https://doi.org/10.3102/0013189X15613981

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

Hanshaw, A. (2024, January 8). Four-day school week faces scrutiny from Missouri legislature, state education board. *The Missouri Independent*. https://missouriindependent.com/2024/01/08/four-day-school-week-faces-scrutiny-from-missouri-legislature-state-education-board/

Ingersoll, R. M., & Tran, H. (2023). Teacher shortages and turnover in rural schools in the US: An organizational analysis. *Educational Administration Quarterly*, *59*(2), 396–431. https://doi.org/10.1177/0013161X231159922

Ketz, J. (2023, February 8). Missouri lawmakers push for mandatory 5-day school week. *Fox4 Kansas City*. https://fox4kc.com/news/missouri-news/missouri-lawmakers-push-for-mandatory-5-day-school-week/

Khalid, A. (2025). Non-monetary incentives in education: investigating the impact of four-day school week on teacher retention & hiring. (Working Paper 2025-09; EDRE Working Papers). https://edre.uark.edu/_resources/pdf/edrewp2025-09.pdf

Kilburn, M. R., Phillips, A., Gomez, C. J., Mariano, L. T., Doss, C. J., Troxel, W. M., Morton, E., & Estes, K. (with RAND Education and Labor (Program), Rand Corporation, & Robert Wood Johnson Foundation). (2021). *Does four equal five? Implementation and outcomes of the four-day school week*. RAND.

Lawson, M. C. (2025). The effect of four-day school week adoption on teacher retention and sorting. (Working Paper 25-1372; EdWorkingPapers). https://edworkingpapers.com/ai25-1372

Morton, E. (2021). Effects of four-day school weeks on school finance and achievement: Evidence from Oklahoma. *Educational Researcher*, *50*(1), 30–40. https://doi.org/10.3102/0013189X20948023

Morton, E. (2023). Effects of 4-day school weeks on older adolescents: Examining impacts of the schedule on academic achievement, attendance, and behavior in high school. *Educational Evaluation and Policy Analysis*, *45*(1), 52–78. https://doi.org/10.3102/01623737221097420

Morton, E., & Dewil, E. (2024). Impacts of four-day school weeks on teacher recruitment and retention and student attendance: Evidence from Colorado (Working Paper 307–0924; CALDER Working Papers). Center for Analysis of Longitudinal Data in Education Research. https://caldercenter.org/publications/impacts-four-day-school-weeks-teacher-recruitment-and-retention-and-student-attendance

Morton, E., Thompson, P. N., & Kuhfeld, M. (2024). A multi-state, student-level analysis of the effects of the four-day school week on student achievement and growth. *Economics of Education Review*, *100*, 102524. https://doi.org/10.1016/j.econedurev.2024.102524

Nowak, A.D., Perrone, F., and Smith, P.S. (2023). How do homeowners, teachers, and students respond to the four-day school week. Working Paper 23-721; EdWorkingPapers). https://edworkingpapers.com/ai23-721.

Podgursky, M., Ehlert, M., Lindsay, J., & Wan, Y. (2016). An examination of the movement of educators within and across three Midwest region states (REL 2017-185). U.S.

Department of Education, Institute of Education Science, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. https://files.eric.ed.gov/fulltext/ED570453.pdf

Reininger, M. (2012). Hometown disadvantage? It depends on where you're from: teachers' location preferences and the implications for staffing schools. *Educational Evaluation and Policy Analysis*, *34*(2), 127–145. https://doi.org/10.3102/0162373711420864

Rhinesmith, E., Anglum, J. C., Park, A., & Burrola, A. (2023). Recruiting and retaining teachers in rural schools: A systematic review of the literature. *Peabody Journal of Education*, *98*(4), 347–363. https://doi.org/10.1080/0161956X.2023.2238491

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, *235*(2), 2218–2244. https://doi.org/10.1016/j.jeconom.2023.03.008

Sant'Anna, P. H. C., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, *219*(1), 101–122. https://doi.org/10.1016/j.jeconom.2020.06.003

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

Thompson, P. N. (2021). Does a day lost equal dollars saved? The effects of four-day school weeks on school district expenditures. *National Tax Journal*, *74*(1), 147–183. https://doi.org/10.1086/712916

Thompson, P. N., Gunter, K., Schuna, Jr., J. M., & Tomayko, E. J. (2021). Are all four-day school weeks created equal? A national assessment of four-day school week policy adoption and implementation. *Education Finance and Policy*, *16*(4), 558–583. https://doi.org/10.1162/edfp_a_00316

Thompson, P. N., Tomayko, E. J., Gunter, K. B., & Schuna, J. (2022). Impacts of the four-day school week on high school achievement and educational engagement. *Education Economics*, *30*(5), 527–539. https://doi.org/10.1080/09645292.2021.2006610

Thompson, P. N., & Ward, J. (2022). Only a matter of time? The role of time in school on four-day school week achievement impacts. *Economics of Education Review*, *86*, 102198. https://doi.org/10.1016/j.econedurev.2021.102198

Turner, J., Finch, K., & Uribe-Zarain, X. (2019). Three midwest rural school districts' first year transition to the four day school week. *The Rural Educator*, *40*(1). https://doi.org/10.35608/ruraled.v40i1.529

Von Hippel, P. T., & Schuetze, B. A. (2025). How not to fool ourselves about heterogeneity of treatment effects. *Advances in Methods and Practices in Psychological Science*, *8*(2), 25152459241304347. https://doi.org/10.1177/25152459241304347

Wooldridge, J. M. (2021). Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences Estimators. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3906345

Figure 1 - Growth of the Four-Day School Week in Missouri

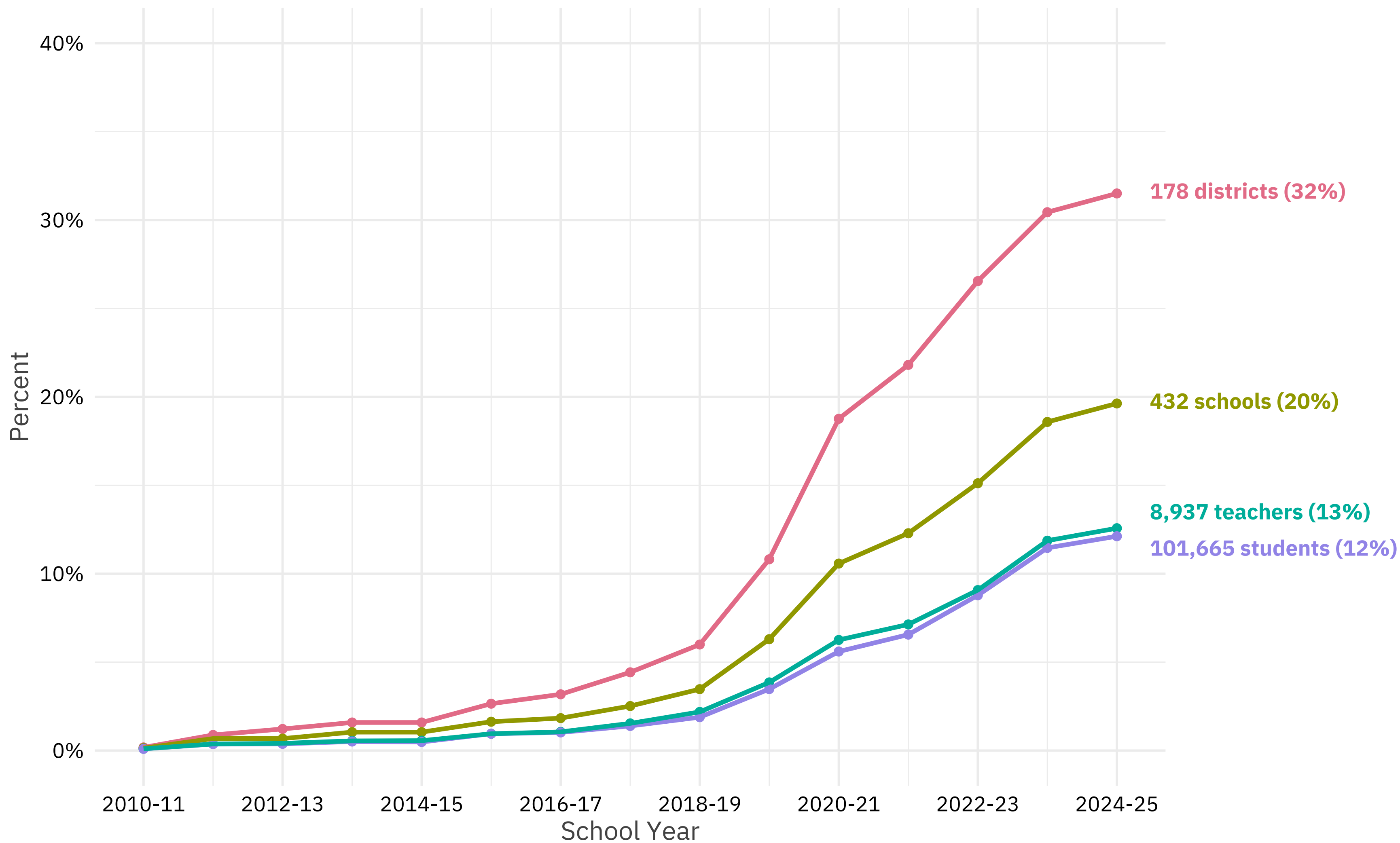
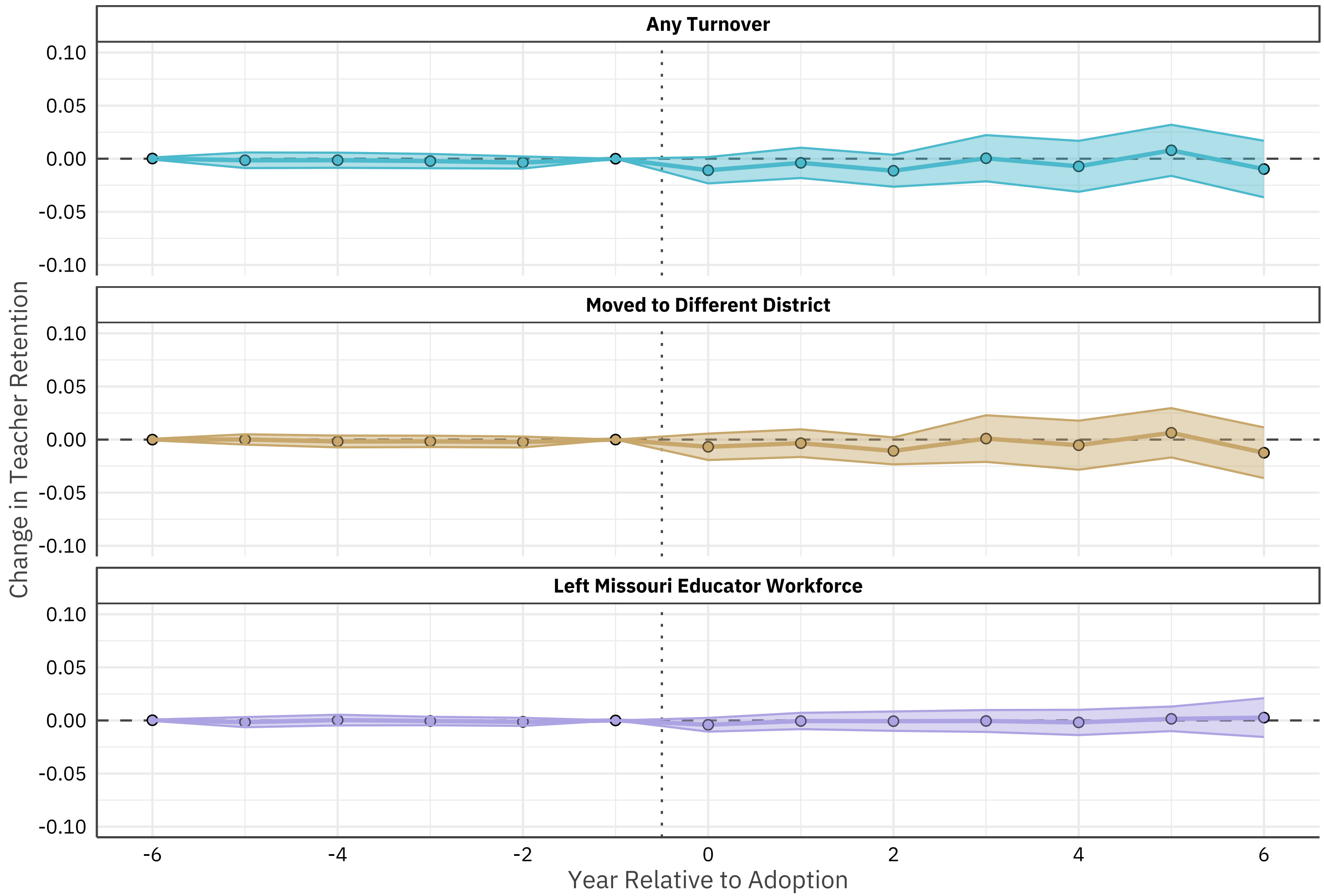
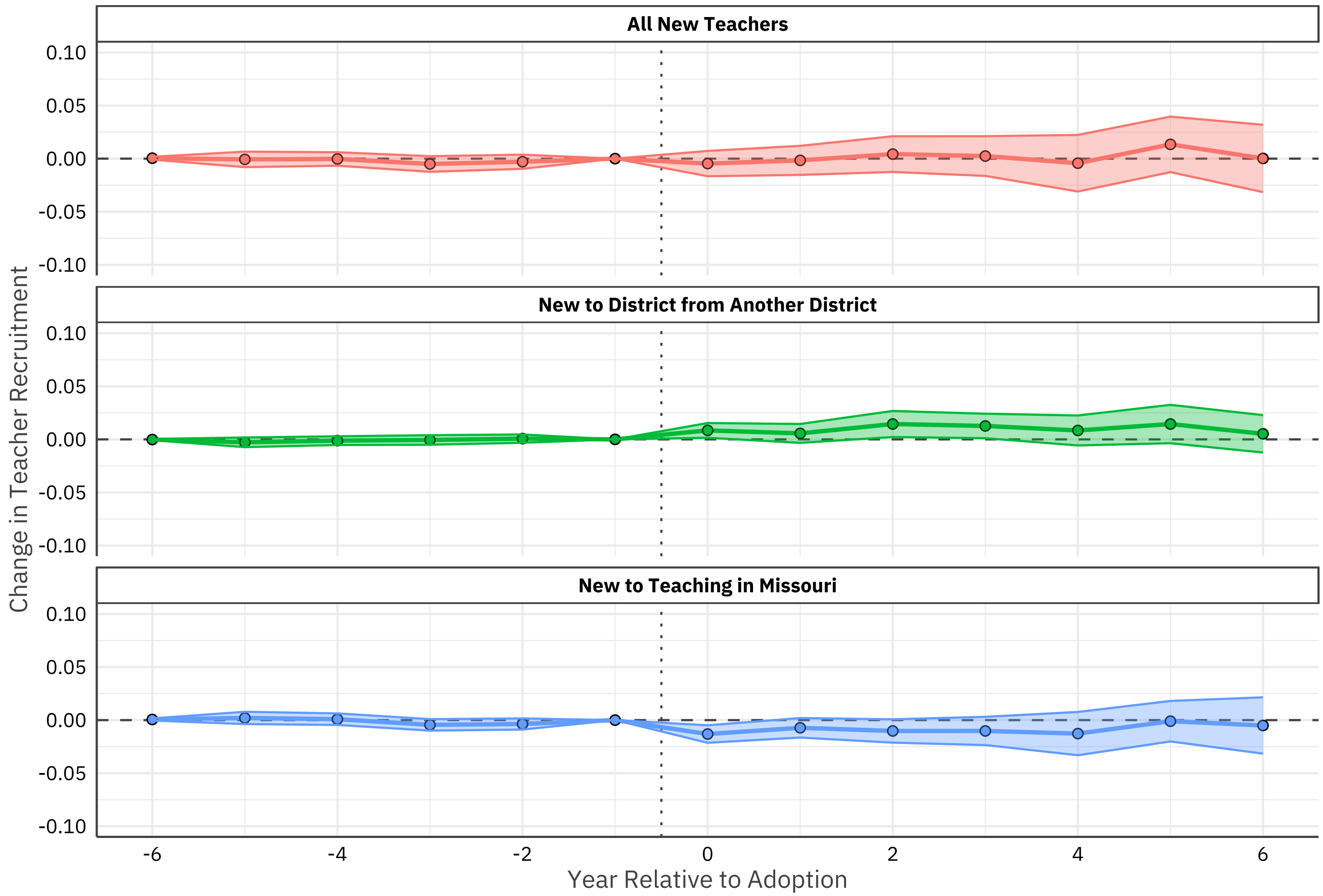


Figure 2 - Event Study Estimates of 4DSW Effects on Teacher Retention (All Teachers)



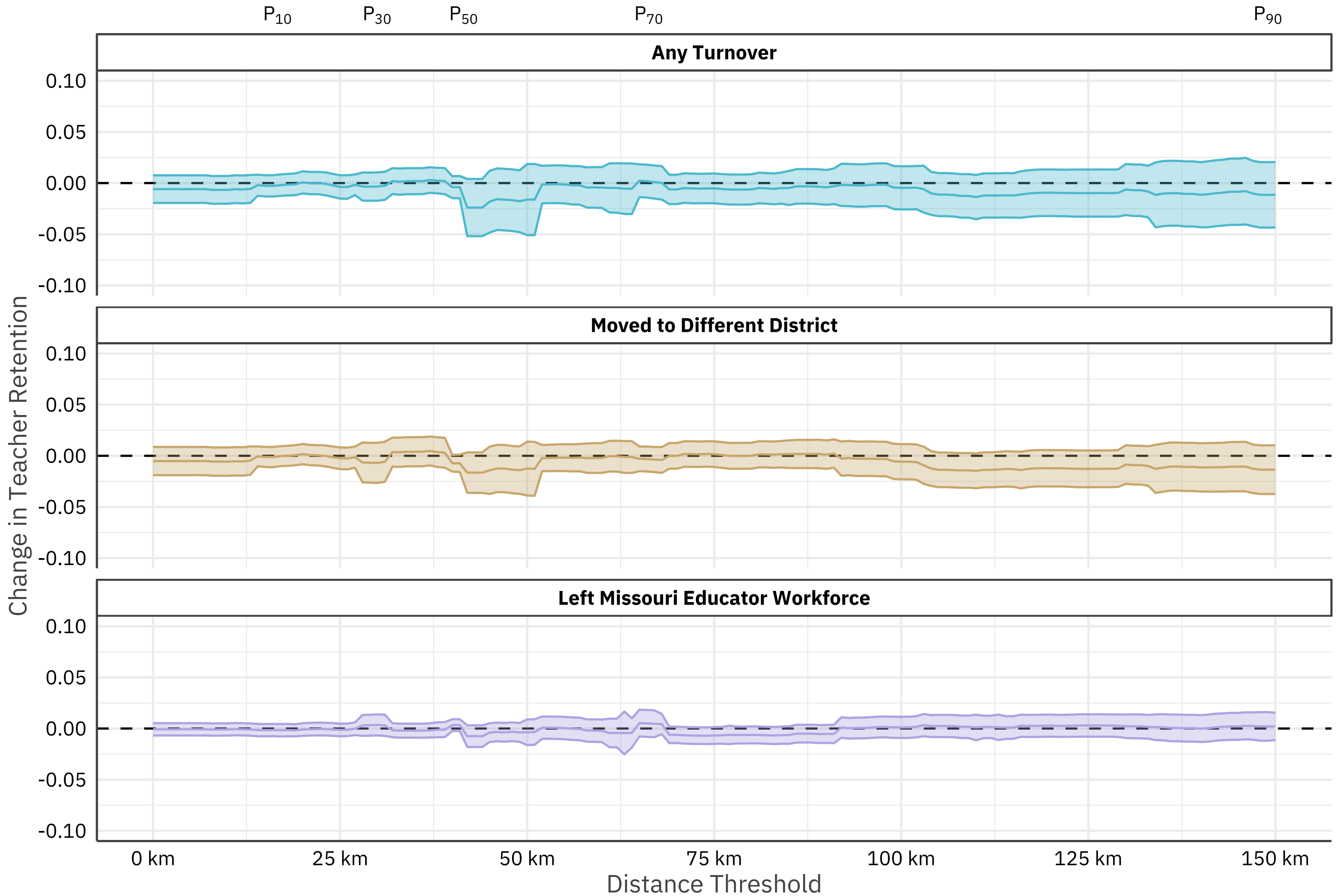
Notes: Event-study analogs to Table 2.

Figure 3 - Event Study Estimates of 4DSW Effects on New Teacher FTE Shares (All Teachers)



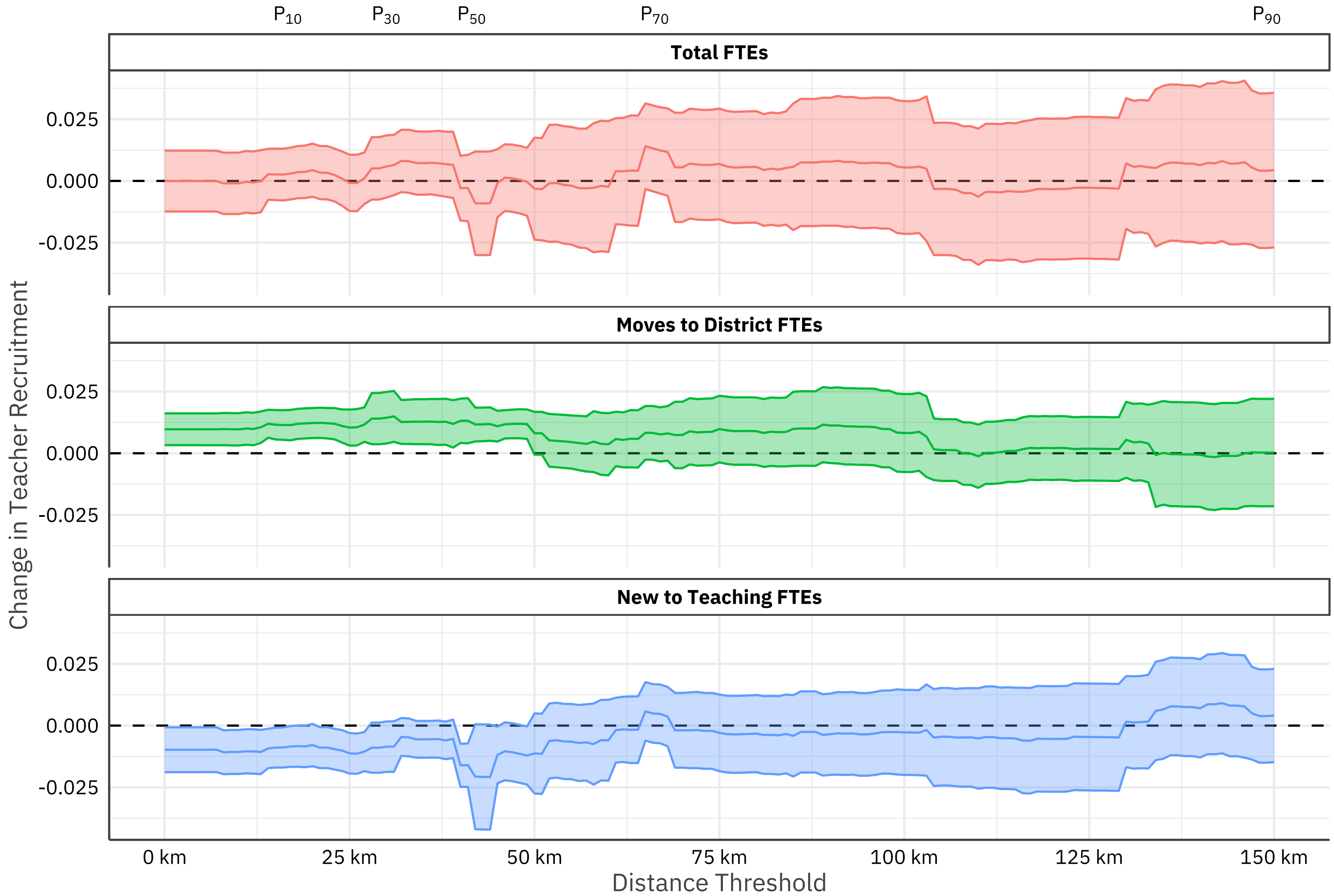
Notes: Event-study analogs to Table 3.

Figure 4 – Tests for Bias in our Primary Retention Estimates from Spatial Spillovers – Teacher Retention



Notes: These graphs show estimates of ω_1 from equation (6) for teacher exits, using d at integer values over the range [0,150] km. The estimates where $d = 0$ replicate our main findings in Table 2, column (1). The upper x-axis indicates specific percentiles in the mover distance distribution.

Figure 5 – Tests for Bias in our Primary Recruitment Estimates from Spatial Spillovers – New Teacher FTE Shares



Notes: These graphs show estimates of ω_1 from equation (6) for the new-teacher FTE shares, using d at integer values over the range [0,150] km. The estimates where d =0 replicate our main findings in Table 3, column (1). The upper x-axis indicates specific percentiles in the mover distance distribution.

Table 1 - Characteristics of School Districts Using 2023-24 Data

	Panel A: Adopters vs. Never-Adopters			*Panel B: Early vs. Late Adopters*		
	Always 5DSW (N=376)	**Ever 4DSW (N=178)**	**Difference in Means**	**Early Adopters (N=106)**	**Late Adopters (N=72)**	**Difference in Means**
District Characteristics						
Enrollment	7,795	2,202	-5,593**	808	3,707	2,899
Number of Schools	14.7	5.5	-9.2**	2.7	8.5	5.8
Teacher FTEs	608.0	164.7	-443.4**	68.0	269.1	201.2
Teacher FTEs (New to District)	79.7	26.7	-53.0**	12.2	42.3	30.1
Teacher FTEs (New to Teaching)	47.9	16.4	-31.5**	5.6	28.1	22.5
Pct. Urban	23.6	0.5	-23.2**	0.9	0.0	-0.9
Pct. Suburban	36.2	10.8	-25.4*	1.0	21.4	20.4
Pct. Town	22.5	9.1	-13.4*	10.0	8.1	-1.9
Pct. Rural	10.9	46.1	35.1**	53.8	37.7	-16.0
Pct. Remote Rural	6.7	33.6	26.9**	34.3	32.7	-1.6
Teacher Characteristics						
Average Salary	$54,909	$44,718	-$10,191**	$43,019	$46,553	$3,533
Average Years Exp.	12.7	12.0	-0.7*	12.1	12.0	-0.1
Pct. MA+	59.2	45.2	-14.0**	46.0	44.4	-1.5
Pct. Any Turnover	14.2	16.9	2.8**	18.5	15.2	-3.3**
Pct. Moved to Different District	5.4	8.4	2.9**	9.6	7.0	-2.6**
Pct. Left Missouri Educator Workforce	8.7	8.5	-0.2	8.9	8.2	-0.7
Pct. New Hires	13.7	16.7	3.1***	17.3	15.9	-1.3
Pct. New to District from Another District	5.9	8.1	2.2***	8.9	6.8	-2.2**
Pct. New to Teaching in Missouri	7.8	8.7	0.9	8.3	9.2	0.9
Student Characteristics						
Pct. Black	16.6	3.5	-13.1**	1.5	5.7	4.2
Pct. Hispanic	8.5	6.7	-1.8	3.9	9.8	5.9
Pct. Other Race/Eth.	8.6	4.9	-3.8**	4.1	5.8	1.7
Pct. White	66.2	84.9	18.7**	90.6	78.8	-11.8
Pct. FRPL	47.9	53.8	5.9	50.8	57.0	6.3

*Note: All descriptive statistics are teacher-weighted district averages, thus reflecting the value or experience of the average Missouri teacher. Late adopters are defined as adopters in 2020-21 or later (though the early versus late differences are similar using different cutoffs). Statistically significant differences are denoted as follows: + p < 0.10, * p < 0.05, ** p < 0.01*

Table 2 - 4DSW Effects on Teacher Turnover

	All Subjects	**In High-Needs**	**In STEM**	**In SPED**
Estimated Effects				
Any Turnover	-0.006	-0.005	-0.013	0.007
	(0.007)	(0.007)	(0.008)	(0.010)
Moved to Different Missouri District	-0.005	-0.004	-0.012	0.006
	(0.007)	(0.006)	(0.008)	(0.007)
Left Missouri Educator Workforce	-0.001	0.000	-0.001	0.001
	(0.003)	(0.004)	(0.007)	(0.008)
Pre-Adoption Means				
Any Turnover	0.126	0.115	0.109	0.110
Moved to Different District	0.073	0.067	0.065	0.064
Left MO Educator Workforce	0.053	0.048	0.044	0.046
Observations (Teacher-Years)	138,633	67,391	41,867	26,032

*Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024). Teacher-level regressions. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Table 3 - 4DSW Effects on New Teacher FTE Shares

	All Subjects	In High-Needs	In STEM	In SPED
Estimated Effects				
All New Hires	0.001	0.000	-0.009	0.022+
	(0.006)	(0.007)	(0.008)	(0.012)
New to District (from Another District)	0.011**	0.010*	0.004	0.029**
	(0.003)	(0.004)	(0.005)	(0.006)
New to Teaching (in Missouri)	-0.010*	-0.010+	-0.012+	-0.007
	(0.004)	(0.006)	(0.007)	(0.009)
Pre-Adoption Means				
All New Hires	0.127	0.121	0.117	0.119
New to District	0.048	0.048	0.048	0.045
New to Teaching	0.079	0.074	0.067	0.075
Observations (District-Years)	2,673	2,673	2,673	2,673

*Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024). District level regressions weighted by teacher FTE. The dependent variable in each regression is the ratio of new-teacher FTE to total teacher FTE; for columns (2)-(4) the total FTE calculation is field specific. Each estimate in the table is from a separate regression. + - $p < 0.10$, * - $p < 0.05$, ** - $p < 0.01$*

Table 4 - 4DSW Effects on Out-of-Field Rates

	High Needs	STEM	SPED
Estimated Effects			
Proportion Out-of-Field	-0.021	-0.031	-0.025
	(0.025)	(0.024)	(0.023)
Observations (District-Years)	2,658	2,657	2,651

*Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024). District level regressions weighted by teacher FTE. The dependent variable in each regression is the ratio of out-of-field FTE to total teacher FTE in the field specified by the column header. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Table 5 - Test for Saturation Effects of 4DSW Calendar Adoption on Teacher Retention

	All Subjects	In High-Needs	In STEM	In SPED
Panel A: Any Turnover				
4DSW	-0.007	-0.005	-0.012	0.009
	(0.007)	(0.008)	(0.010)	(0.014)
4DSW × Prevalence	0.010	0.002	-0.008	-0.041
	(0.050)	(0.061)	(0.072)	(0.102)
Panel B: Moved to Different Missouri District				
4DSW	-0.006	-0.002	-0.011	0.013
	(0.006)	(0.007)	(0.008)	(0.011)
4DSW × Prevalence	0.009	-0.030	-0.012	-0.099
	(0.042)	(0.050)	(0.057)	(0.080)
Panel C: Left Missouri Educator Workforce				
4DSW	-0.001	-0.003	-0.001	-0.003
	(0.004)	(0.005)	(0.007)	(0.008)
4DSW × Prevalence	0.001	0.032	0.004	0.058
	(0.029)	(0.037)	(0.049)	(0.056)
Observations (Teacher-Years)	138,633	67,391	41,867	26,032

*Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024). Teacher-level regressions. Note the interaction with the prevalence of other 4DSW districts indicates the (infeasible) effect of going from 0 to 100 percent of FTE in the commuting zone in a 4DSW district. The standard deviation of the CZ-level share of FTE in a 4DSW district is roughly 0.095. Each pair of estimates in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Table 6 – Test for Saturation Effects of 4DSW Calendar Adoption on Teacher Recruitment

	All Subjects	In High-Needs	In STEM	In SPED
Panel A: All New Hires				
4DSW	-0.002	-0.004	-0.010	0.015
	(0.008)	(0.009)	(0.010)	(0.014)
4DSW × Prevalence	0.029	0.051	0.048	0.076
	(0.054)	(0.065)	(0.070)	(0.130)
Panel A: New to District (from another MO district)				
4DSW	0.009**	0.008+	0.003	0.025***
	(0.003)	(0.005)	(0.006)	(0.007)
4DSW × Prevalence	0.008	0.012	-0.002	0.051
	(0.027)	(0.037)	(0.045)	(0.078)
Panel A: New to Teaching in Missouri				
4DSW	-0.010+	-0.013+	-0.013+	-0.009
	(0.006)	(0.007)	(0.007)	(0.011)
4DSW × Prevalence	0.021	0.039	0.050	0.025
	(0.041)	(0.050)	(0.052)	(0.084)
Observations (District -Years)	2,673	2,673	2,673	2,673

*Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024). Teacher-level regressions. Note the interaction with the prevalence of other 4DSW districts indicates the (infeasible) effect of going from 0 to 100 percent of FTE in the commuting zone in a 4DSW district. The standard deviation of the CZ-level share of FTE in a 4DSW district is roughly 0.095. Each pair of estimates in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

APPENDIX A – ROBUSTNESS TO ALTERNATIVE ESTIMATORS

This appendix shows that our findings are substantively robust to a number of modifications to our estimation approach. First, in Table A.1, we estimate our retention models using district-aggregated data rather than teacher-level data, to align the level of aggregation with our analysis of teacher recruitment. Given that our identifying variation is at the district level already, this has a negligible effect on our results as anticipated. Next, in Table A.2, we confirm our retention findings are qualitatively similar if we add standard time-varying control variables to our main models. We include controls for teacher experience, degree, race, gender, and workload as a full-time equivalent share (FTE), as well as district enrollment, the racial/ethnic composition of students, and the share of students enrolled for free- or reduced-priced lunch. Conceptually, we prefer to exclude these variables because including them requires imposing assumptions about the effect of treatment on their evolution, but we show these estimates for completeness.

Next, in Tables A.3, A.4, and A.5, we replicate our retention analysis using three different staggered difference-in-differences approaches (Borusyak et al., 2024; Callaway & Sant'Anna, 2021; Wooldridge, 2021). While in some specifications we find significant impacts of the 4DSW on STEM teacher retention, most of our estimates are insignificant, and all are small, aligning with our findings presented in the main text. Tables A.6, A.7, A.8, and A.9 show analogous results that correspond to new-teacher recruitment. Again, these results are broadly consistent with the results in the main text.

Table A.1 - 4DSW Effects on Teacher Turnover (District Level)

	All Subjects	In High-Needs	In STEM	In SPED
Any Turnover	-0.006	-0.005	-0.013	0.007
	(0.007)	(0.007)	(0.008)	(0.010)
Moved to Different Missouri District	-0.005	-0.004	-0.012+	0.006
	(0.007)	(0.006)	(0.007)	(0.007)
Left Missouri Educator Workforce	-0.001	0.000	-0.001	0.001
	(0.003)	(0.004)	(0.005)	(0.006)
Observations (District-Years)	2,674	2,658	2,652	2,609

*Note: Estimated via two-stage Difference-in-Differences (Gardner et al., 2024). District-level regressions weighted by teacher FTEs (and subject-specific teacher FTEs in columns 2-4). Each estimate in the table is from a separate regression. + - $p < 0.10$, * - $p < 0.05$, ** - $p < 0.01$*

Table A.2 - 4DSW Effects on Teacher Turnover (Gardner et al. Estimator w/ Controls)

	All Subjects	In High-Needs	In STEM	In SPED
Any Turnover	-0.005	-0.006	-0.015+	0.005
	(0.007)	(0.007)	(0.008)	(0.013)
Moved to Different Missouri District	-0.004	-0.004	-0.015*	0.007
	(0.007)	(0.007)	(0.008)	(0.010)
Left Missouri Educator Workforce	-0.001	-0.001	0.000	-0.001
	(0.003)	(0.005)	(0.007)	(0.008)
Observations (Teacher-Years)	138,181	67,157	41,713	25,945

*Note: Estimated via two-stage Difference-in-Differences (Gardner et al., 2024) with controls for teacher (experience, degree level, FTE, race, and gender) and district-level (enrollment, share of non-white students, and percent of students receiving free- or reduced-priced lunch) characteristics. Teacher-level regressions. Each estimate in the table is from a separate regression. + - $p < 0.10$, * - $p < 0.05$, ** - $p < 0.01$*

Table A.3 - 4DSW Effects on Teacher Turnover (DiD Imputation)

	All Subjects	In High-Needs	In STEM	In SPED
Any Turnover	-0.006	-0.005	-0.013+	0.007
	(0.007)	(0.006)	(0.007)	(0.012)
Moved to Different Missouri District	-0.005	-0.004	-0.012	0.006
	(0.007)	(0.007)	(0.008)	(0.008)
Left Missouri Educator Workforce	-0.001	0.000	-0.001	0.001
	(0.003)	(0.004)	(0.006)	(0.007)
Observations (Teacher-Years)	138,633	67,391	41,867	26,032

*Note: Estimated via difference-in-differences imputation (Borusyak, Jaravel, & Spiess, 2024). Teacher-level regressions. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Table A.4 - 4DSW Effects on Teacher Turnover (ETWFE)

	All Subjects	In High-Needs	In STEM	In SPED
Any Turnover	-0.007	-0.006	-0.011	0.004
	(0.007)	(0.006)	(0.007)	(0.012)
Moved to Different Missouri District	-0.006	-0.005	-0.010	0.004
	(0.007)	(0.007)	(0.008)	(0.008)
Left Missouri Educator Workforce	-0.001	-0.001	-0.001	0.000
	(0.003)	(0.004)	(0.007)	(0.007)
Observations (Teacher-Years)	138,633	67,391	41,867	26,032

*Note: Estimated via extended two-way fixed effects (Wooldridge, 2021). Teacher-level regressions. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Table A.5 - 4DSW Effects on Teacher Turnover (Callaway & Sant'Anna)

	All Subjects	In High-Needs	In STEM	In SPED
Any Turnover	-0.02	-0.026	-0.028	-0.020
	(0.025)	(0.048)	(0.062)	(0.086)
Moved to Different Missouri District	-0.016	-0.018	-0.015	-0.018
	(0.011)	(0.040)	(0.071)	(0.022)
Left Missouri Educator Workforce	-0.004	-0.008	-0.012	-0.002
	(0.015)	(0.030)	(0.012)	(0.099)
Observations (Teacher-Year)	138,633	67,391	41,867	26,032

*Note: Estimated via Callaway & Sant'Anna (2021) nonparametric estimator with group-time ATTs aggregated as a weighted average where weights correspond to the number of teachers treated in each group-time cell. Teacher-level regressions. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Table A.6 - 4DSW Effects on New Teacher FTE Shares (Gardner et al. Estimator w/ Controls)

	All Subjects	In High-Needs	In STEM	In SPED
All New Teachers	0.004	0.002	-0.005	0.021+
	(0.006)	(0.007)	(0.009)	(0.012)
New to District (from Another District)	0.010**	0.010*	0.004	0.028**
	(0.003)	(0.004)	(0.007)	(0.006)
New to Teaching (in Missouri)	-0.006	-0.008	-0.009	-0.007
	(0.004)	(0.006)	(0.007)	(0.009)
Observations (District-Years)	2,673	2,673	2,673	2,673

Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024) with controls for district-level (enrollment, share of non-white students, and percent of students receiving free- or reduced-priced lunch) characteristics. District level regressions weighted by teacher FTE. Each estimate in the table is from a separate regression.
*+ - p < 0.10, * - p < 0.05, ** - p < 0.01*

Table A.7 - 4DSW Effects on New Teacher FTE Shares (DiD Imputation)

	All Subjects	In High-Needs	In STEM	In SPED
All New Teachers	0.001	0.000	-0.014+	0.025*
	(0.004)	(0.005)	(0.007)	(0.010)
New to District	0.010**	0.010*	0.002	0.029**
(from Another District)	(0.002)	(0.004)	(0.007)	(0.005)
New to Teaching	-0.010**	-0.010**	-0.015*	-0.004
(in Missouri)	(0.003)	(0.004)	(0.008)	(0.007)
Observations (District-Years)	2,673	2,673	2,673	2,673

*Note: Estimated via difference-in-differences imputation (Borusyak, Jaravel, & Spiess, 2024). District level regressions weighted by teacher FTE. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Table A.8 - 4DSW Effects on New Teacher FTE Shares (ETWFE)

	All Subjects	In High-Needs	In STEM	In SPED
All New Teachers	-0.001	-0.003	-0.016+	0.024*
	(0.006)	(0.007)	(0.009)	(0.012)
New to District	0.010**	0.009+	0.001	0.030**
(from Another District)	(0.003)	(0.005)	(0.008)	(0.006)
New to Teaching	-0.011*	-0.011*	-0.017*	-0.006
(in Missouri)	(0.004)	(0.006)	(0.008)	(0.009)
Observations (District-Years)	2,673	2,673	2,673	2,673

*Note: Estimated via extended two-way fixed effects (Wooldridge, 2021). District level regressions weighted by teacher FTE. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Table A.9 - 4DSW Effects on New Teacher FTE Shares (Callaway & Sant'Anna)

	All Subjects	In High-Needs	In STEM	In SPED
All New Teachers	-0.013	-0.009	-0.019	0.025
	(0.020)	(0.024)	(0.022)	(0.036)
New to District (from Another District)	0.000	-0.003	-0.014	0.018
	(0.006)	(0.011)	(0.013)	(0.011)
New to Teaching (in Missouri)	-0.013	-0.006	-0.005	0.007
	(0.015)	(0.014)	(0.013)	(0.036)
Observations (District-Years)	2,673	2,673	2,673	2,673

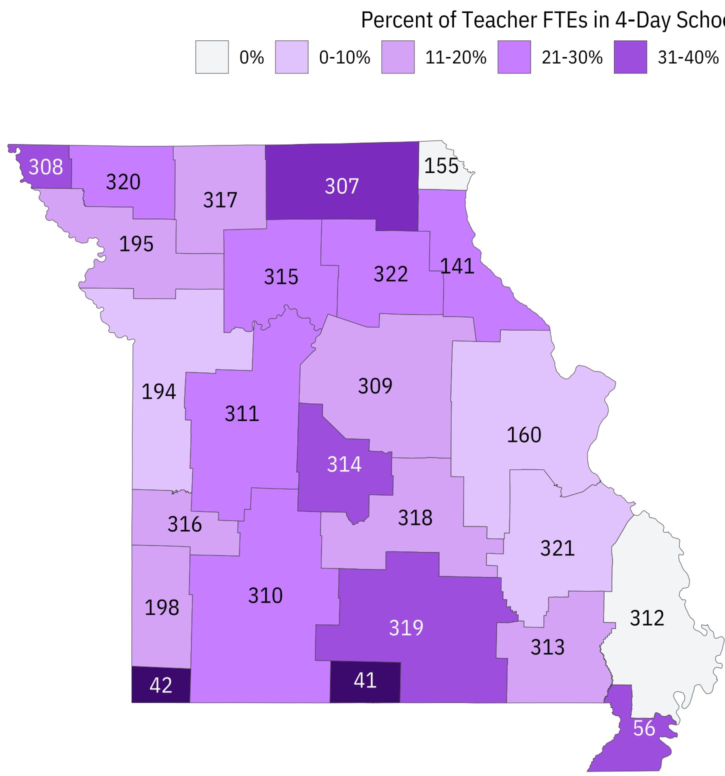
*Note: Estimated via Callaway & Sant'Anna (2021) nonparametric estimator with group-time ATTs aggregated as a weighted average where weights correspond to the number of teacher FTE treated in each group-time cell. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Appendix B – Saturation Effects of 4DSW Calendars

In this appendix we provide additional information about our tests for the moderating effect of local market saturation (corresponding to equation (7) in the main text). To measure market saturation we use commuting zones, which are contiguous sets of counties identified based on commuting flows as reported in the American Community Survey (Fowler, 2024). Missouri contains 25 commuting zones and as of the 2024-25 school year, 23 had at least one 4DSW district. Figure B.1 illustrates prevalence of the 4DSW in Missouri's commuting zones in 2024-25. The variation we use to identify saturation effects is both spatial and temporal, and illustrated in greater detail in Figure B.2. The spatial variation can be seen across rows within columns, while the temporal variation is across columns within rows. Darker colors indicate the 4DSW is more prevalent.

Figures B.3 and B.4 report event-study analogs of the non-interacted term in equation (7) and corroborate the results presented in the main text.

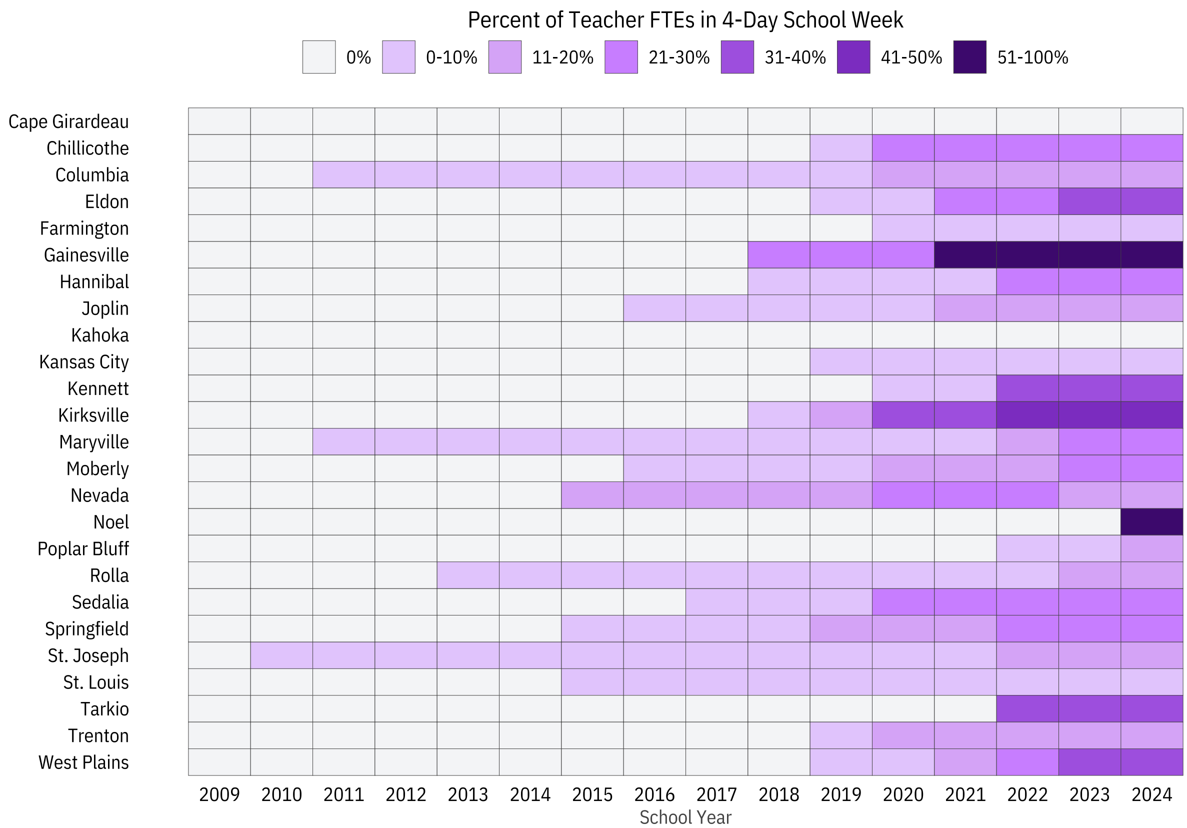
Figure B.1 – Prevalence of 4DSW Calendars by Commuting Zone (2024-25 School Year)



ID	Zone Name	ID	Zone Name
41	Gainesville	311	Sedalia
42	Noel	312	Cape Girardeau
56	Kennett	313	Poplar Bluff
141	Hannibal	314	Eldon
155	Kahoka	315	Chillicothe
160	St. Louis	316	Nevada
194	Kansas City	317	Trenton
195	St. Joseph	318	Rolla
198	Joplin	319	West Plains
307	Kirksville	320	Maryville
308	Tarkio	321	Farmington
309	Columbia	322	Moberly
310	Springfield		

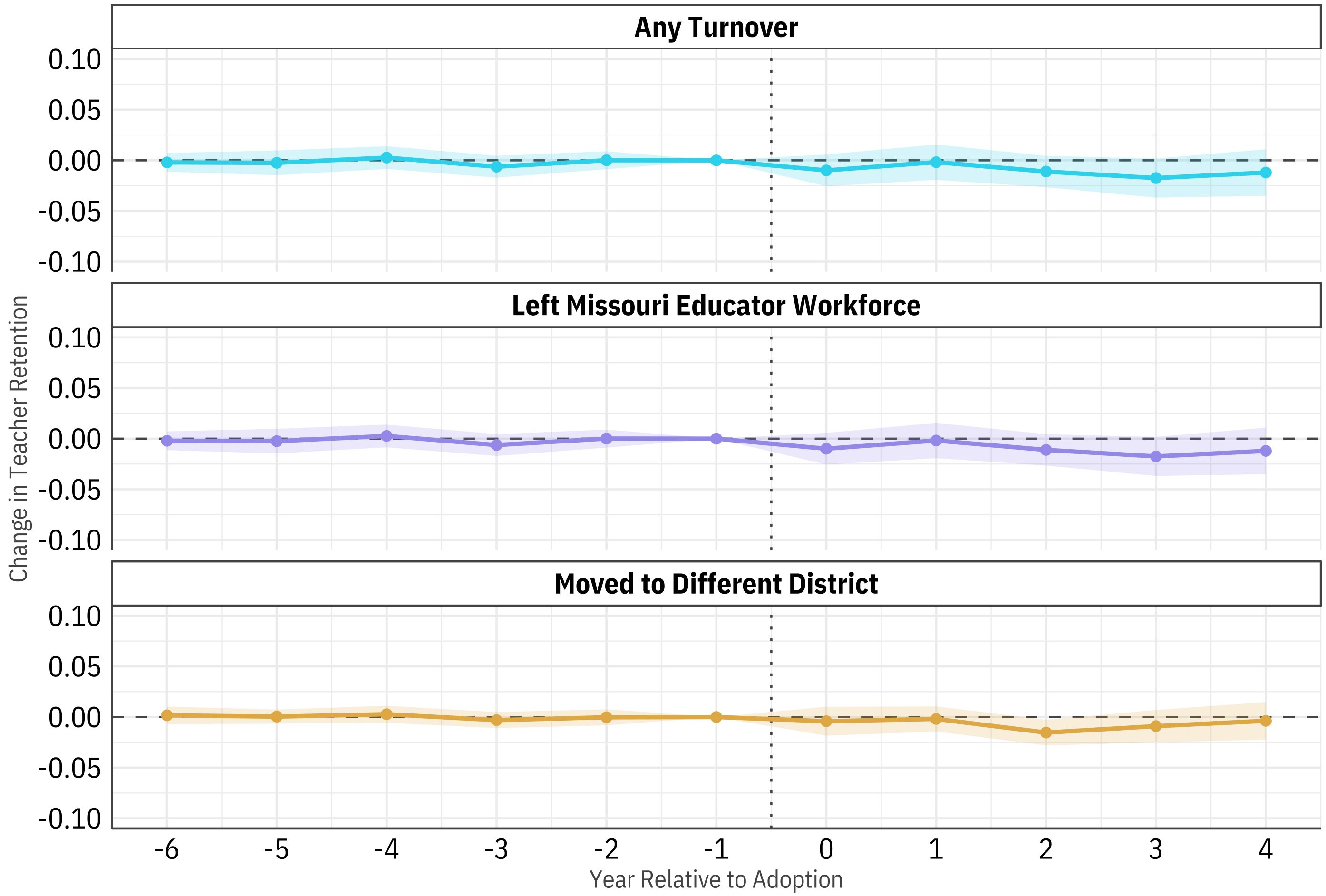
Notes: This map shows the prevalence of four-day school week calendars by commuting zone (CZ) as of the 2024-25 school year. The prevalence of four-day school weeks by commuting zones (CZs) calculated as the share of teacher FTEs in districts using a 4DSW calendar during the 2024-25 school year. While CZs are not officially named, we label each CZ by the largest city it contains here.

Figure B.2 – Variation in 4DSW Prevalence by Commuting Zone and School Year



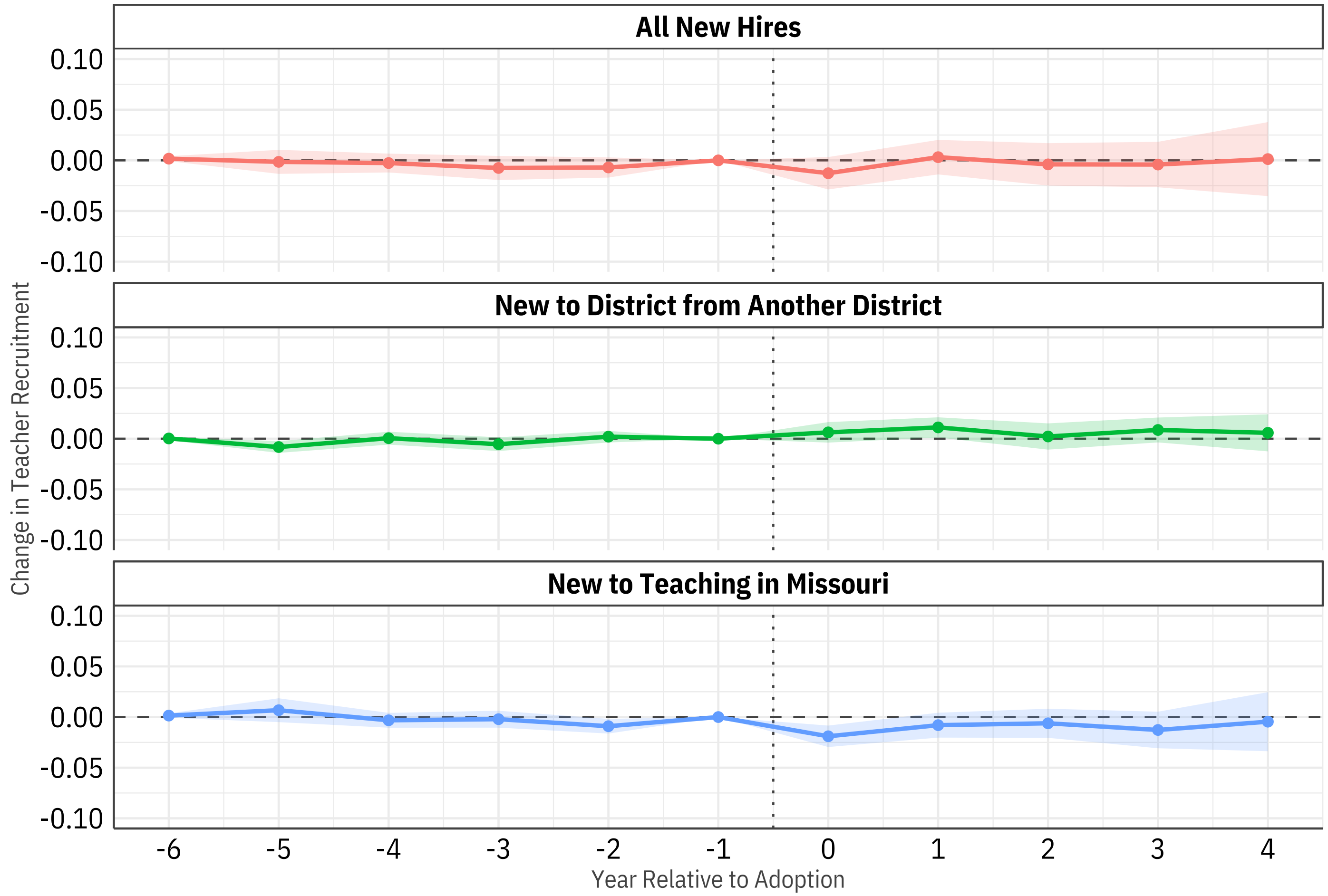
Notes: The prevalence of four-day school weeks by commuting zones (CZs) calculated as the share of teacher FTEs in 4DSW districts for each commuting zone and school year. While CZs are not officially named, we label each CZ by the largest city it contains here.

Figure B.3 – Event Study Estimates of 4DSW Effects on Teacher Retention Net of Saturation Effects (All Teachers)



Notes: Event-study analogs to Table 5. This plot shows the interaction of the 4DSW treatment indicator with relative time indicators and captures the estimated effect of adopting a 4DSW calendar when the prevalence of 4DSWs within the commuting zone is equal to zero.

Figure B.4 – Event Study Estimates of 4DSW Effects on Teacher Recruitment Net of Saturation Effects (All Teachers)



Notes: Event-study analogs to Table 6. This plot shows the interaction of the 4DSW treatment indicator with relative time indicators and captures the estimated effect of adopting a 4DSW calendar when the prevalence of 4DSWs within the commuting zone is equal to zero.

APPENDIX C – ROBUSTNESS TO CHOICE OF COMPARISON GROUP & EFFECTS ON OTHER OUTCOMES

There is strong selection into adoption of the 4DSW in Missouri (and elsewhere). As we show in Table 1, Missouri districts that ever choose to adopt the 4DSW differ in many observable ways from districts that do not. Though the identifying assumptions of our difference-in-differences models do not require "level equivalence" along any dimension—only parallel trends, which can exist with level differences—the large differences in Table 1 nonetheless raise concerns about using never-treated districts as a control group for 4DSW adopters. Accordingly, we follow the modern difference-in-differences literature and use not-yet-treated observations as our control group to improve comparability.

To probe this decision further, in this appendix we consider two alternative options for expanding the control group: (1) a control group that consists of not-yet-treated districts plus rural never-treated districts, and (2) a control group that consists of not-yet-treated districts plus all never-treated districts. To explore the viability of these alternatives, we estimate a series of event studies using six district characteristics as outcomes: student enrollment, share of non-white students, share of students enrolled for free and reduced-price lunch, average teacher experience, share of teachers with a master's degree or more, and residualized teacher salary (conditional on a cubic of experience and teacher education). We estimate the event studies using our preferred control group and the two alternatives. These tests go beyond looking at pre-trends for our outcome of interest, providing broader evidence on the comparability between 4DSW adopters and the control groups.

The event study estimates are reported in figures C.1 – C.6. Each event study is estimated using the Gardner et al. (2024) 2-stage difference-in-differences estimator and weighted by teacher FTEs. Looking across these figures, there are differential pre-trends for student enrollment and the share of students who receive free- or reduced-priced lunch when we use the largest control group inclusive of all never-treated districts. Both of these trends could be linked to school staffing needs—and especially student enrollment—which suggests the control group inclusive of all never-treated districts may not provide a good counterfactual.

Alternatively, when we expand the control group to include only rural never-treated districts, there is no indication of concerning pre-trends for any of the six variables we consider. We prefer the standard not-yet-treated control group, but Appendix Tables C.1 and C.2 show that our findings are very similar to what we report in the main text if we use this alternative control group instead.

Table C.1 - 4DSW Effects on Teacher Turnover (Using Alternate Control Group Including Never Treated Rural Districts)

	All Subjects	In High-Needs	In STEM	In SPED
Any Turnover	-0.006	-0.003	-0.002	0.001
	(0.004)	(0.005)	(0.006)	(0.008)
Moved to Different Missouri District	-0.006+	-0.001	-0.002	0.007
	(0.003)	(0.004)	(0.005)	(0.007)
Left Missouri Educator Workforce	-0.001	-0.002	0.000	-0.007
	(0.003)	(0.003)	(0.004)	(0.004)
Observations (Teacher-Years)	289,432	142,870	87,643	55,996

*Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024). Teacher-level regressions. Each estimate in the table is from a separate regression. + - $p < 0.10$, * - $p < 0.05$, ** - $p < 0.01$*

Table C.2 - 4DSW Effects on New Teacher FTE Shares (Using Alternate Control Group Including Never Treated Rural Districts).

	All Subjects	In High-Needs	In STEM	In SPED
All New Hires	0.002	0.004	-0.003	0.020*
	(0.005)	(0.006)	(0.006)	(0.009)
New to District (from Another District)	0.008**	0.008*	0.004	0.018**
	(0.003)	(0.003)	(0.004)	(0.005)
New to Teaching (in Missouri)	-0.006	-0.003+	-0.007	0.002
	(0.004)	(0.004)	(0.004)	(0.007)
Observations (District-Years)	5,718	5,697	5,673	5,595

*Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024). District level regressions weighted by teacher FTE. The dependent variable in each regression is the ratio of new-teacher FTE to total teacher FTE; for columns (2)-(4) the total FTE calculation is field specific. Each estimate in the table is from a separate regression. + - $p < 0.10$, * - $p < 0.05$, ** - $p < 0.01$*

Figure C.1 – Changes in Student Enrollment by Potential Comparison Group

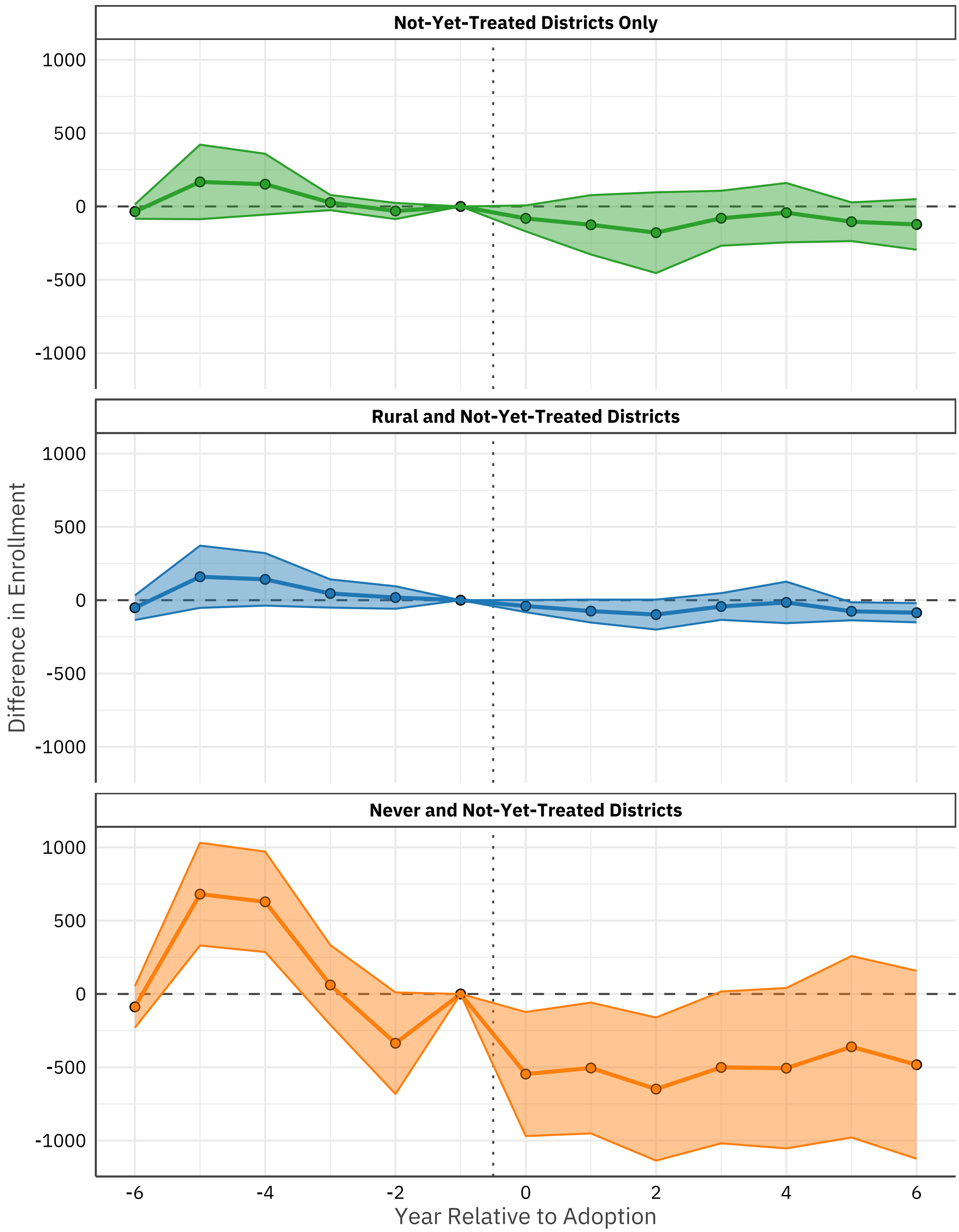


Figure C.2 – Changes in Share of Non-White Students by Potential Comparison Group

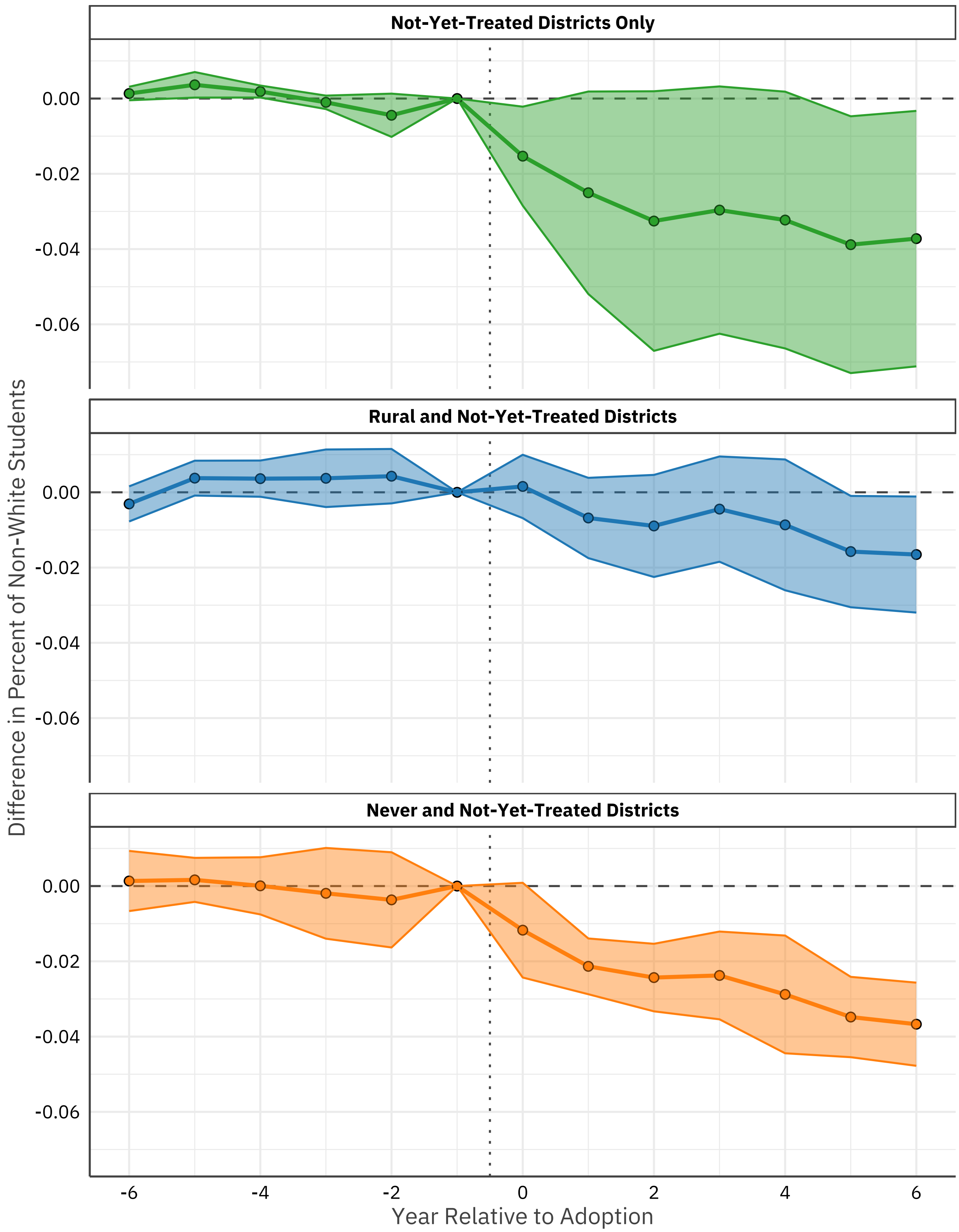


Figure C.3 – Changes in Share of FRPL Students by Potential Comparison Group

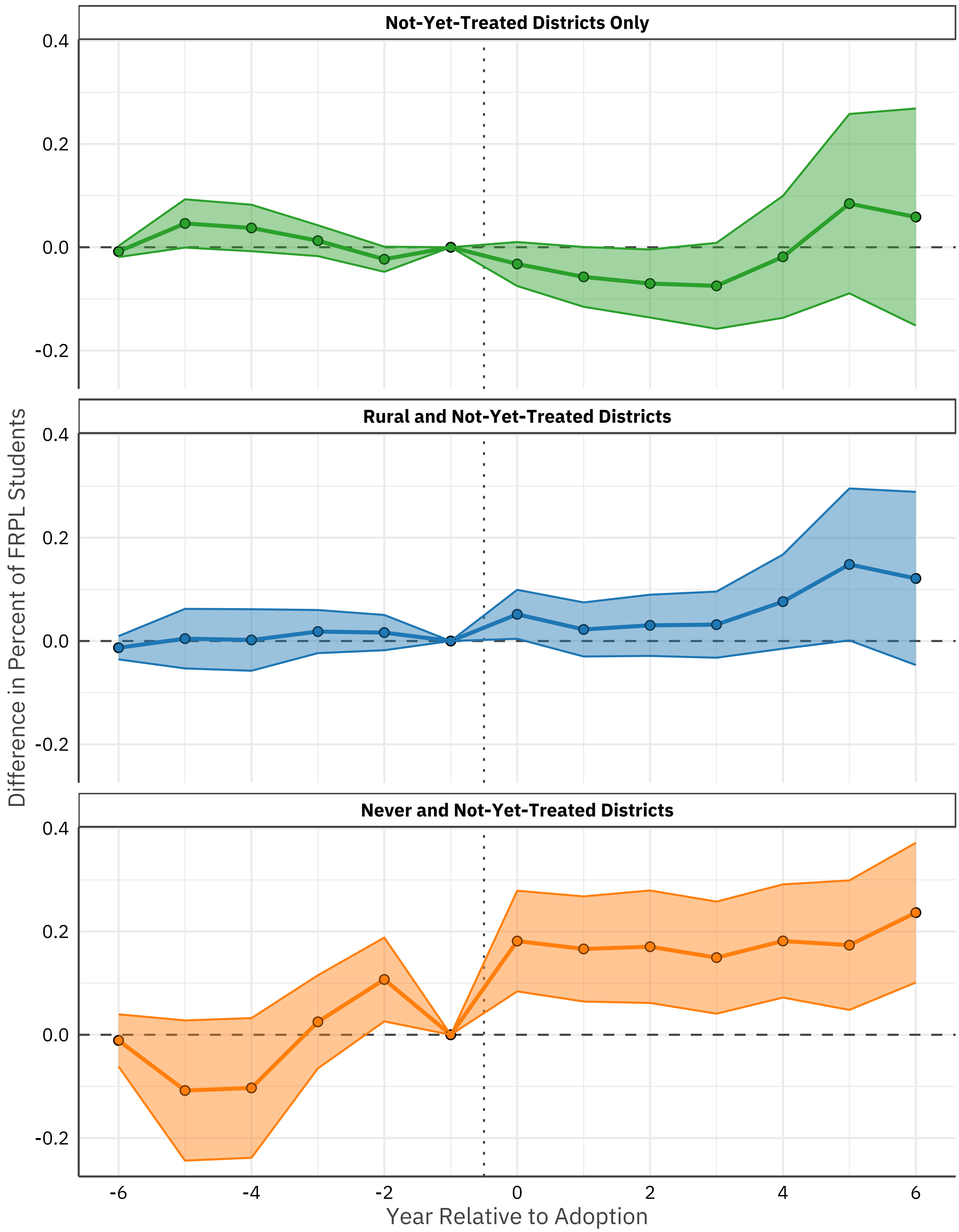


Figure C.4 – Changes in Average Teacher Experience by Potential Comparison Group

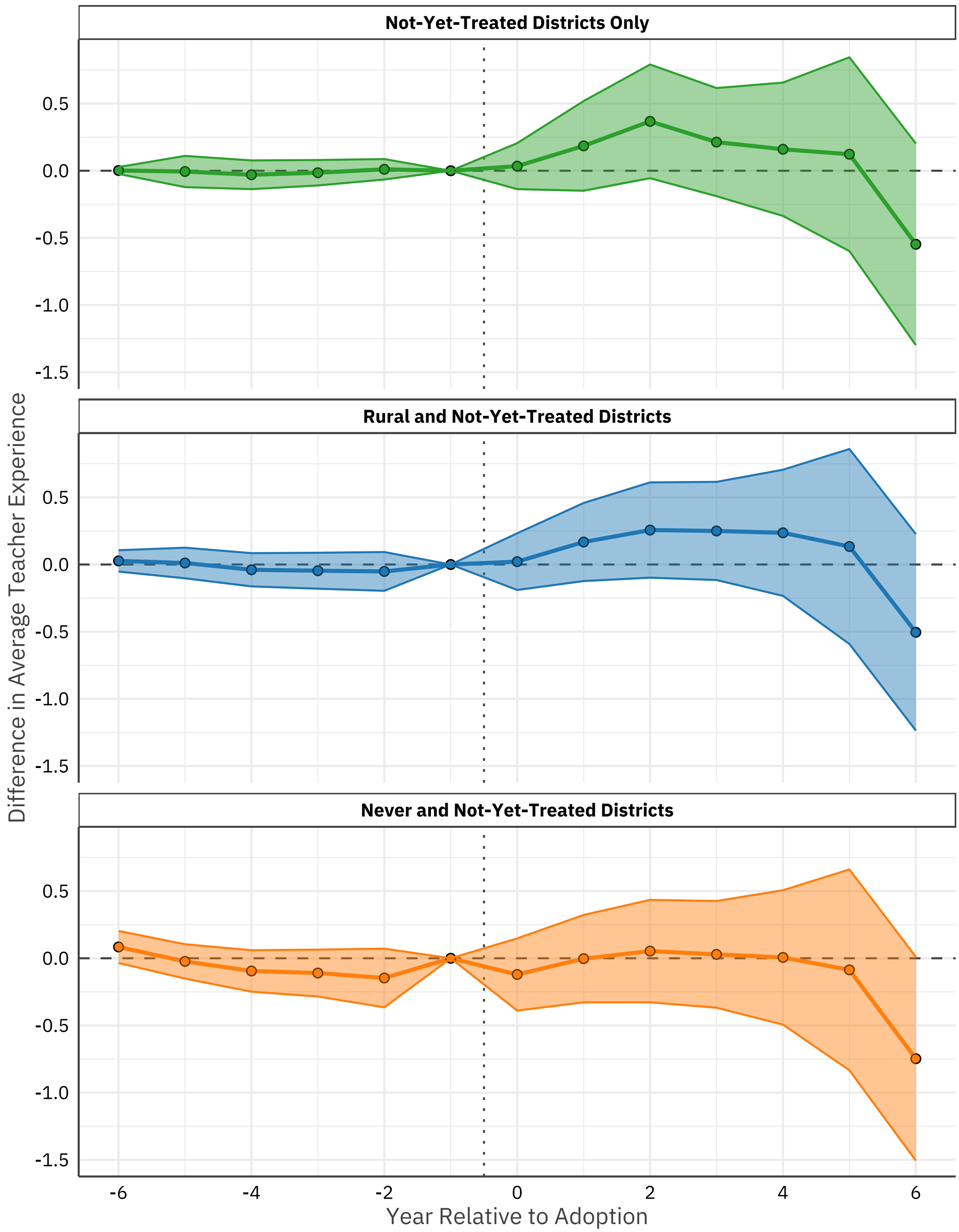


Figure C.5 – Changes in Share of Teachers with MA+ Degree by Potential Comparison Group

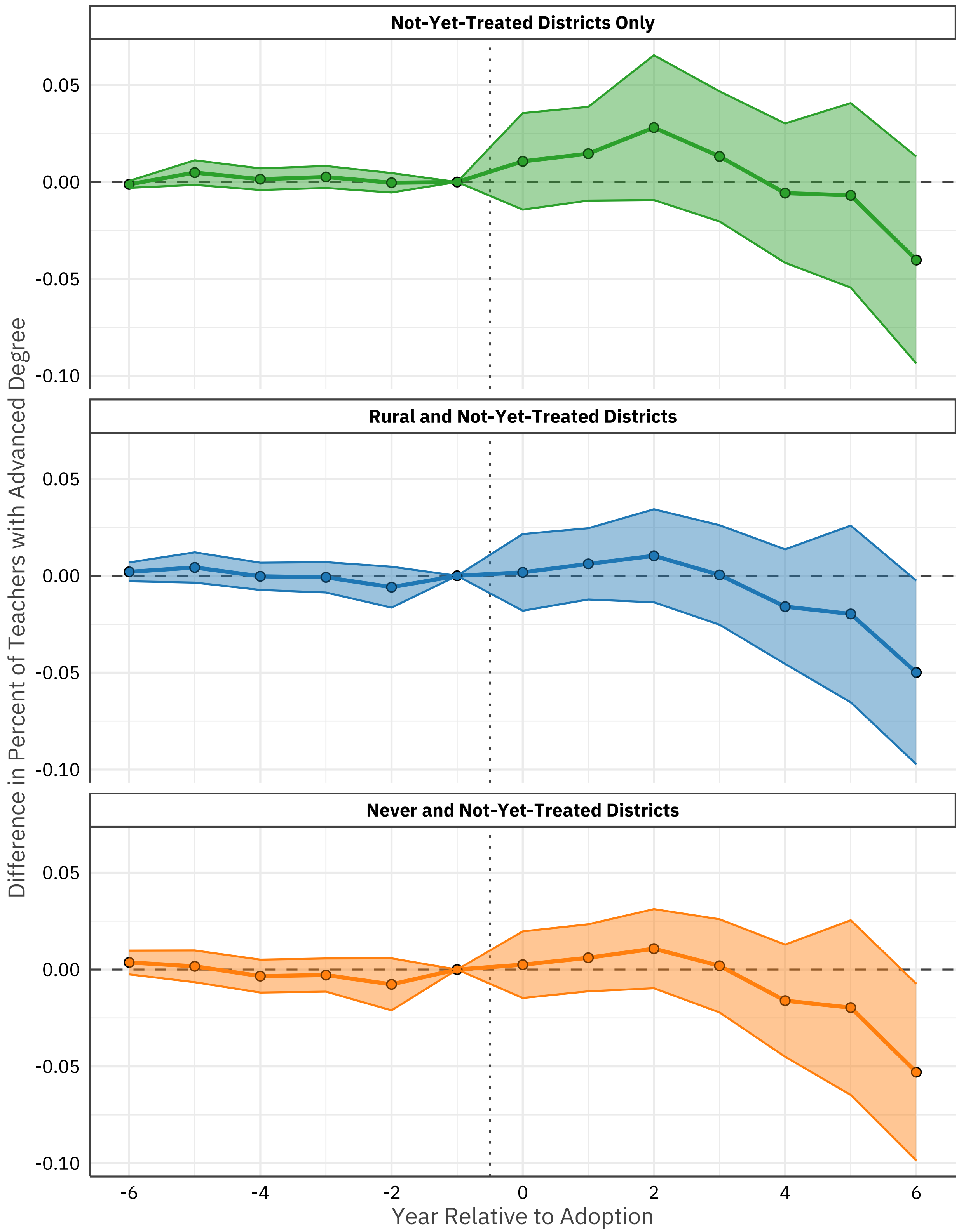
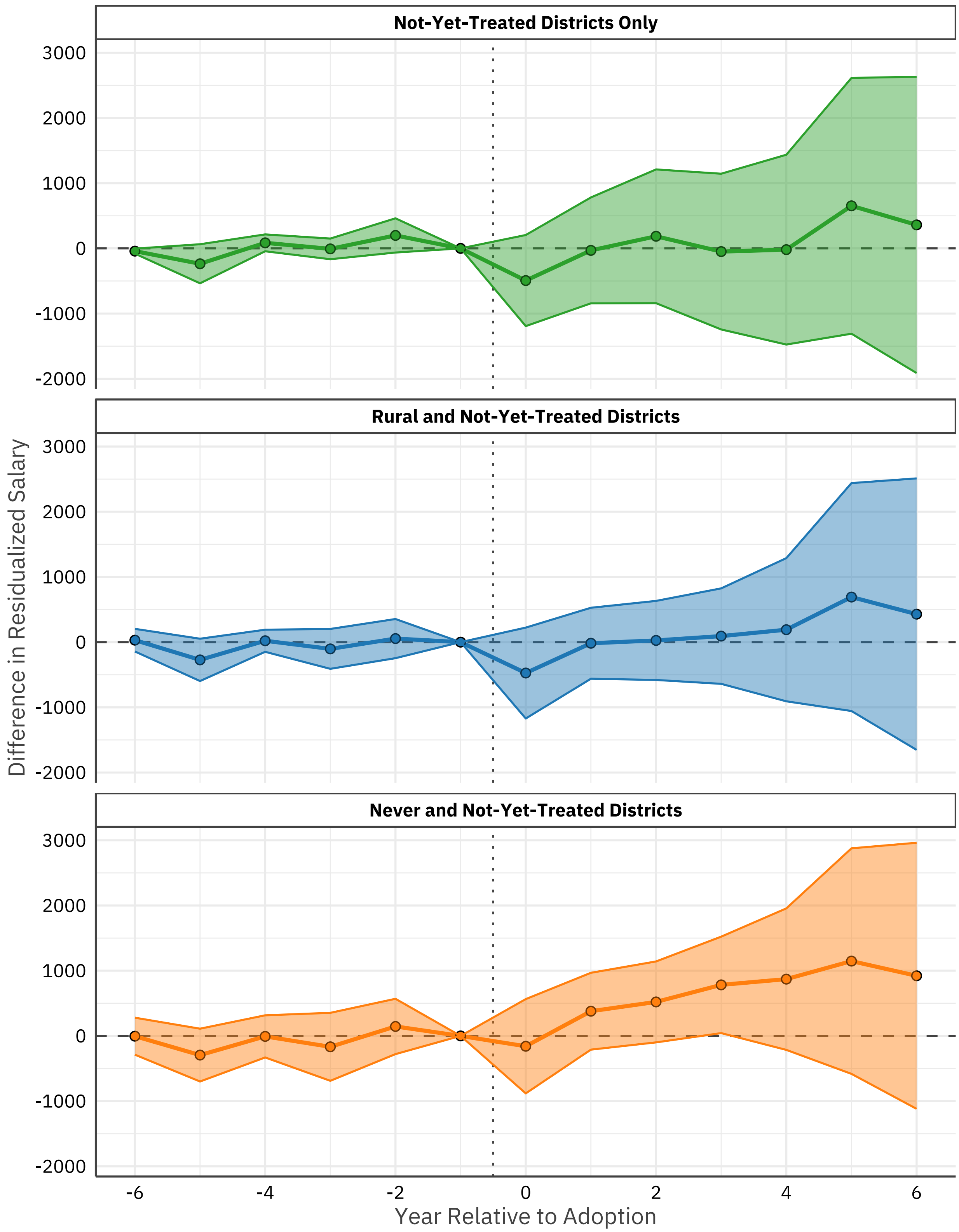


Figure C.6 – Changes in Residualized Teacher Salary by Potential Comparison Group



APPENDIX D – EFFECTS FOR INTERVIEWED DISTRICTS

In this appendix we test for the possibility that districts in the interview sample have had more success with the 4DSW than the average district. If this were true, it could explain the contrast between the optimism expressed about the 4DSW in the interviews and our mostly null quantitative results. To test for this, we replicate our primary analysis but drop all districts from the treatment sample that were not represented in the interviews. That is, we estimate the effects of the 4DSW only in interviewed districts, relative to the same counterfactual of not-yet-treated districts used in the main text. With the caveat that our estimates are somewhat less precise when we restrict the treatment group in this way, there is no evidence to suggest the 4DSW has been more effective in districts represented in the interview sample. We report results analogous to our main results in Tables D.1 and D.2, with event studies shown in Figures D.1 and D.2.

Table D.1 - 4DSW Effects on Teacher Turnover (Excluding Treated Observations from Non-Interviewed districts)

	All Subjects	In High-Needs	In STEM	In SPED
Estimated Effects				
Any Turnover	-0.007	-0.012	-0.022+	-0.003
	(0.010)	(0.010)	(0.012)	(0.017)
Moved to Different Missouri District	-0.007	-0.007	-0.013	-0.006
	(0.010)	(0.011)	(0.011)	(0.016)
Left Missouri Educator Workforce	0.000	-0.004	-0.009	0.002
	(0.006)	(0.009)	(0.010)	(0.012)
Pre-Adoption Means				
Any Turnover	0.119	0.124	0.125	0.116
Moved to Different District	0.065	0.068	0.068	0.066
Left MO Educator Workforce	0.054	0.056	0.056	0.049
Observations (Teacher-Years)	110,753	53,124	32,480	20,813

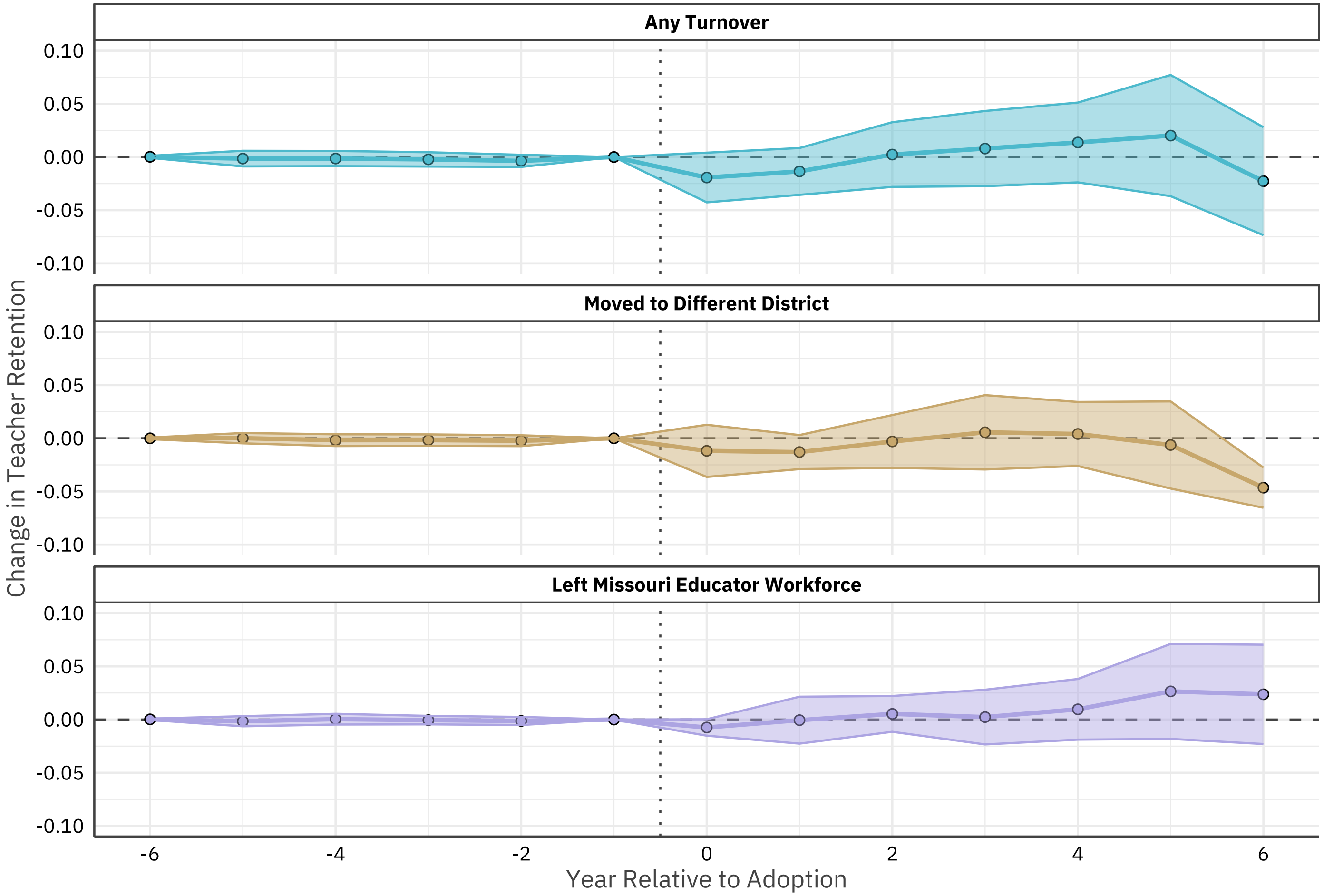
*Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024). Teacher-level regressions. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Table D.2 - 4DSW Effects on New Teacher FTE Shares (Excluding Treated Observations from Non-Interviewed districts)

	All Subjects	**In High-Needs**	**In STEM**	**In SPED**
Estimated Effects				
All New Hires	-0.008	-0.009	0.003	-0.010
	(0.011)	(0.013)	(0.013)	(0.017)
New to District (from Another District)	0.007	0.006	0.007	0.021**
	(0.006)	(0.006)	(0.008)	(0.010)
New to Teaching (in Missouri)	-0.014+	-0.015	-0.004	-0.031***
	(0.008)	(0.012)	(0.012)	(0.012)
Pre-Adoption Means				
All New Hires	0.126	0.128	0.126	0.125
New to District	0.045	0.048	0.050	0.043
New to Teaching	0.081	0.079	0.075	0.080
Observations (District-Years)	2,033	2,033	2,033	2,033

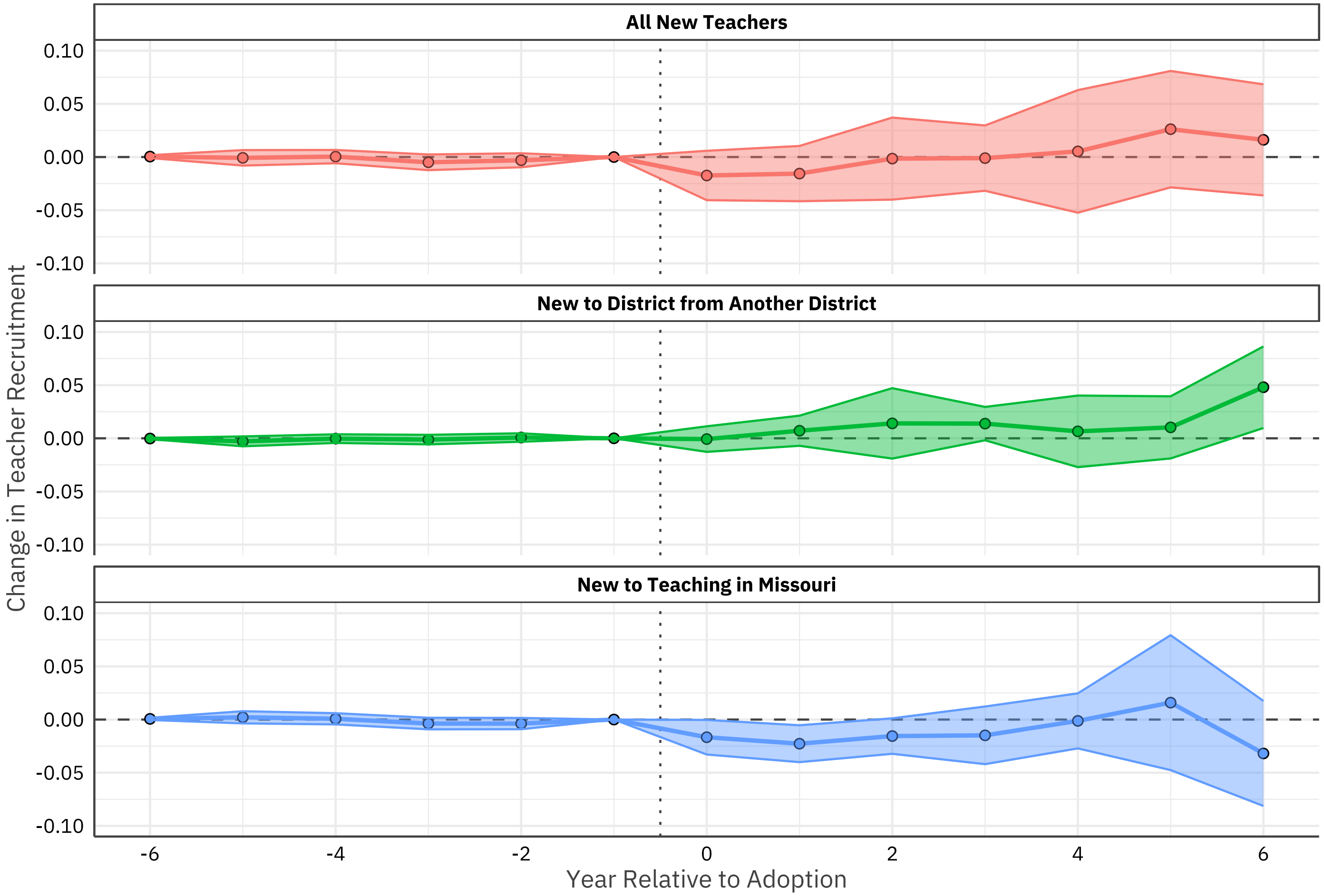
*Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024). District level regressions weighted by teacher FTE. The dependent variable in each regression is the ratio of new-teacher FTE to total teacher FTE; for columns (2)-(4) the total FTE calculation is field specific. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Figure D.1 - Event Study Estimates of 4DSW Effects on Teacher Retention
(All Teachers; Excluding observations from non-interviewed districts in treatment group)



Notes: Event-study analogs to Table D.1.

Figure D.2 - Event Study Estimates of 4DSW Effects on Teacher Recruitment (All Teachers; Excluding observations from non-interviewed districts in treatment group)



Notes: Event-study analogs to Table D.2.

APPENDIX E –EFFECTS IN PRE-COVID PERIOD

Our analysis in the main text is inclusive of the COVID pandemic and its associated disruption. Our models broadly control for the average effects of the pandemic with year fixed effects. However, it is possible that the pandemic dampened the effects of the 4DSW specifically—for instance, by changing mobility and hiring practices for reasons unrelated to workforce amenities. To test for this possibility, we replicate our main models but restrict the sample to pre-COVID years (specifically, we only include retention and new hires through the 2019-20 school year). The results omitting the pandemic period are very similar to our results reported in the main text, indicating our findings are not unduly influenced by the pandemic. We report results analogous to our main results in Tables E.1 and E.2, with event studies shown in Figures E.1 and E.2.

Table E.1 - 4DSW Effects on Teacher Turnover (Excluding the COVID Pandemic Period)

	All Subjects	In High-Needs	In STEM	In SPED
Panel A: Any Turnover				
4DSW	-0.007	0.001	0.004	0.007
	(0.006)	(0.009)	(0.010)	(0.015)
Panel B: Moved to Different District				
4DSW	-0.008+	0.003	0.002	0.010
	(0.005)	(0.007)	(0.007)	(0.010)
Panel C: Left Missouri Educator Workforce				
4DSW	0.002	-0.002	0.002	-0.003
	(0.005)	(0.006)	(0.006)	(0.011)
Observations (Teacher-Years)	101,715	48,773	29,993	18,787

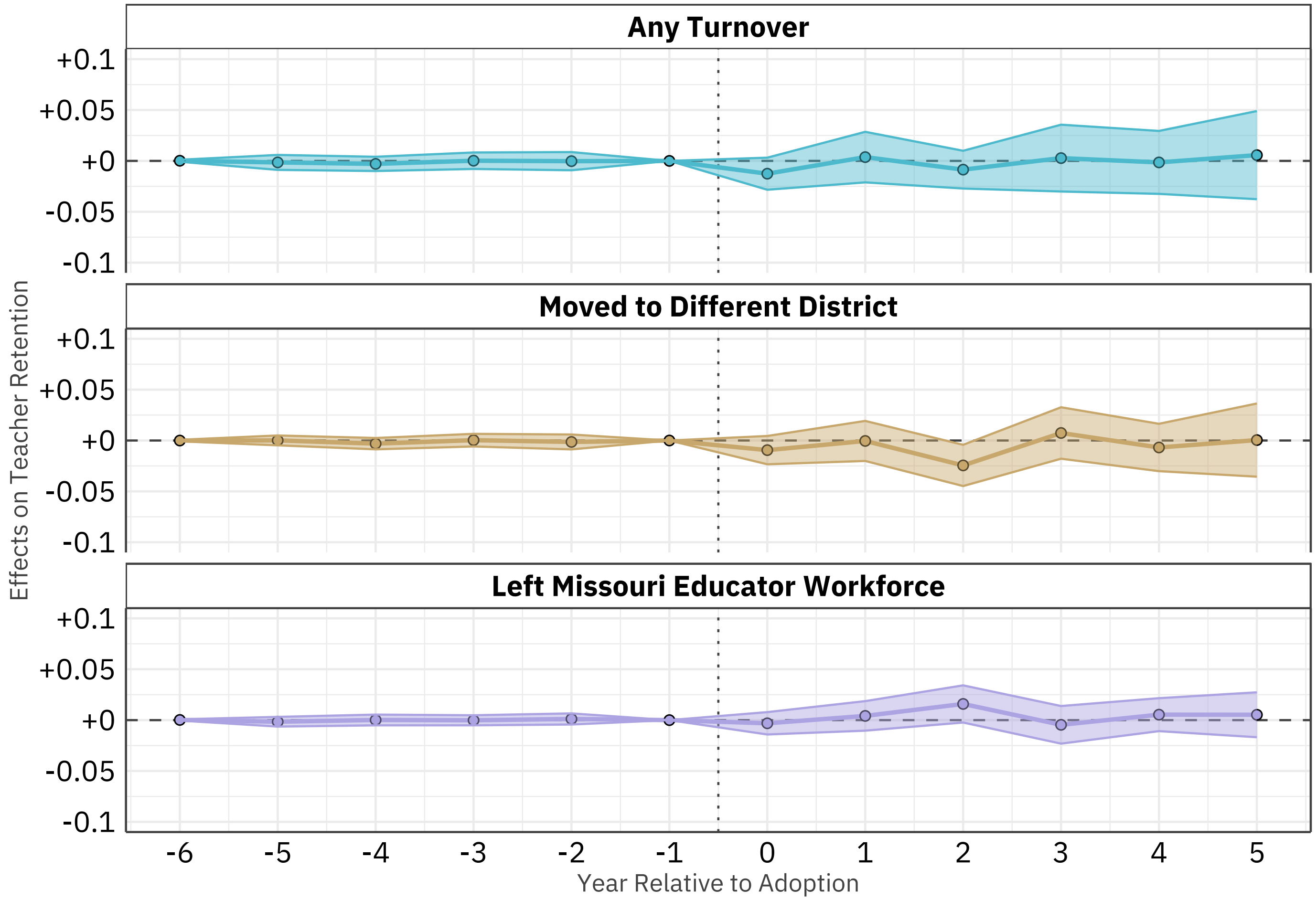
*Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024). Teacher-level regressions. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Table E.2 - 4DSW Effects on New Teacher FTE Shares (Excluding the COVID Pandemic Period)

	All Subjects	In High-Needs	In STEM	In SPED
Panel A: All New Hires				
4DSW	0.007	-0.001	-0.011	0.015
	(0.010)	(0.012)	(0.016)	(0.019)
Panel A: New to District (from another MO district)				
4DSW	0.010**	0.006	-0.004	0.026***
	(0.005)	(0.006)	(0.007)	(0.008)
Panel A: New to Teaching in Missouri				
4DSW	-0.003	-0.007	-0.007	-0.011
	(0.008)	(0.009)	(0.012)	(0.018)
Observations (District -Years)	1,773	1,773	1,773	1,773

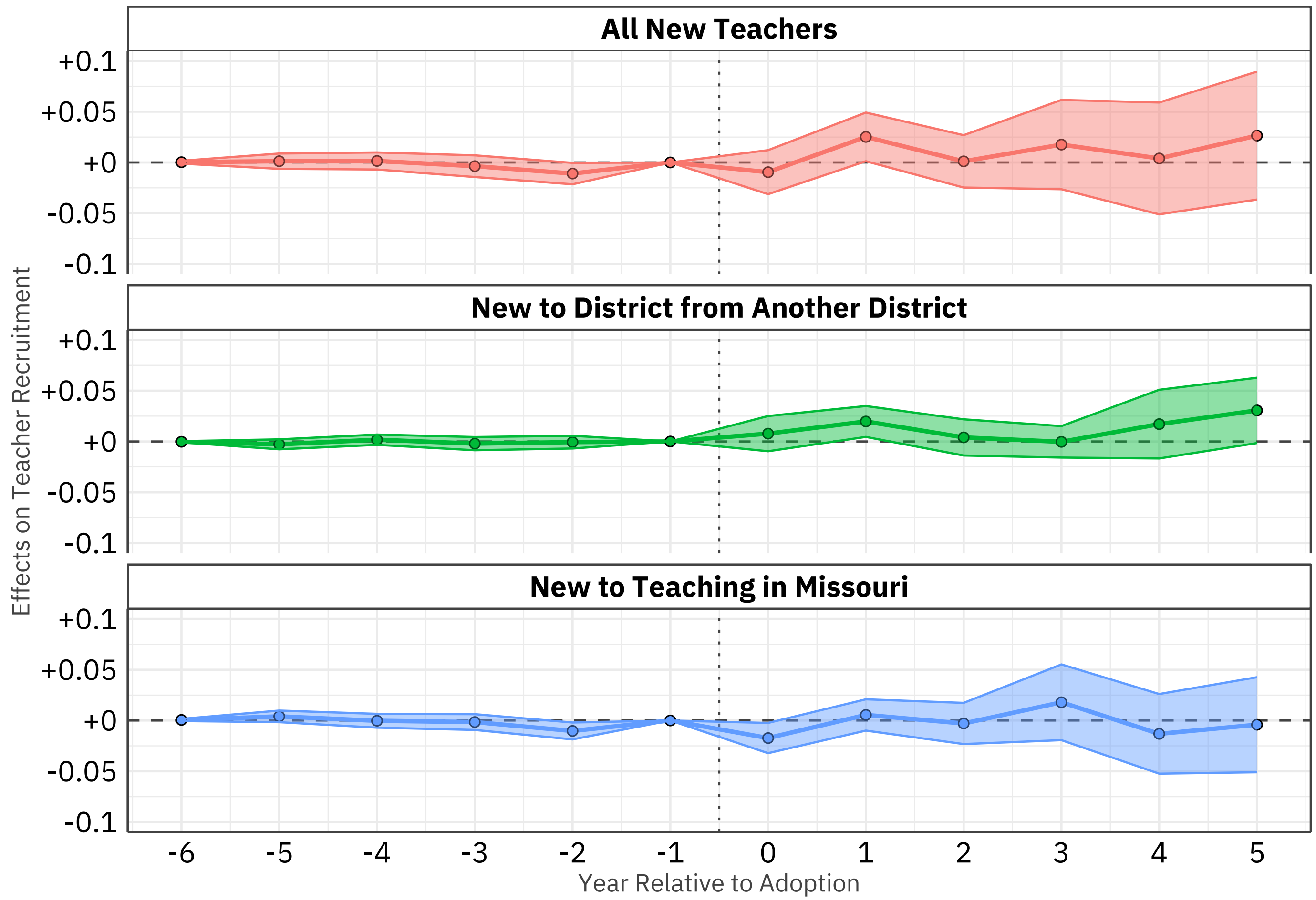
*Note: Estimated via two-stage difference-in-differences (Gardner et al., 2024). Teacher-level regressions. Each estimate in the table is from a separate regression. + - p < 0.10, * - p < 0.05, ** - p < 0.01*

Figure E.1 - Event Study Estimates of 4DSW Effects on Teacher Retention
(All Teachers; Excluding the COVID Pandemic Period)



Notes: Event-study analogs to Table E.1.

Figure E.2 - Event Study Estimates of 4DSW Effects on Teacher Recruitment
(All Teachers; Excluding the COVID Pandemic Period)



Notes: Event-study analogs to Table E.2.