Evaluating the Impact of Voucher Programs on Incumbent Students in Participating Private Schools

Brianna Felegi* R. Joseph Waddington†

bfelegi@vt.edu rwadding@nd.edu

September 28, 2026

Abstract

In this paper, we examine the impacts of the Indiana Choice Scholarship Program on incumbent students in voucher-participating private schools using a mixed methods approach. We find that higher shares of voucher students are associated with declines in test scores, particularly in the early in the program and among students at the tails of the achievement distribution. Contemporaneous interview evidence suggests these effects reflect increased student academic heterogeneity, instructional adaptation, and changes in peer dynamics. Our findings suggest that early voucher program implementation can generate short-run disruptions, with implications for how new and expanding programs are designed and evaluated.

*Department of Economics, Virginia Polytechnic Institute and State University

†Institute for Educational Initiatives, University of Notre Dame

This paper was supported by Notre Dame's Center for Research on Educational Opportunity (CREO) and the Institute of Educational Initiatives. We are grateful to the Indiana Department of Education for providing access to state administrative records and for supporting independent analyses. We are also grateful to Roberto Peñaloza, who helped organize the data. We are especially grateful to Chloe Gibbs and David Figlio for their helpful comments. We also thank participants at the 2024 Association for Public Policy Analysis and Management Fall Research Conference, the 2025 Association for Education Finance and Policy Annual Conference, the 2025 Mid-South Education Policy Workshop, the 2025 Midwest Economics of Education Conference, and participants in seminars at the University of Rochester, Virginia Tech and the College of William & Mary for their insightful feedback as this paper has evolved. All opinions expressed in this paper represent those of the authors and not necessarily the institutions with which they are affiliated. All errors in this paper are solely the responsibility of the authors. The authors report no conflicts of interest.

I Introduction

Private school choice programs, including vouchers, tax credit-scholarships, and education savings accounts, are becoming increasingly popular policy tools to increase access to schooling options in the United States. Since the inception of the first voucher program in 1990, there has been a large increase in the number of state funded programs and a growth in the share of students eligible to participate. As a result, a large literature examines the impacts of these programs on students using vouchers (Rouse, 1998; Mayer et al., 2002; Howell et al., 2002; Wolf et al., 2010; Witte et al., 2014; Chingos and Peterson, 2015; Abdulkadiroğlu et al., 2018; Waddington and Berends, 2018) and the competitive effects on students remaining in the public school system (Hoxby, 2003; Figlio and Rouse, 2006; Chakrabarti, 2008; Chiang, 2009; Greene and Marsh, 2009; Chakrabarti, 2013; Rouse et al., 2013; Figlio and Hart, 2014; Chingos et al., 2025). However, much less attention has been focused on how voucher programs impact participating private schools and their existing students.

Although a growing body of research has examined how private schools respond to voucher programs, most of this work focuses on changes in enrollment, finances, school inputs, and overall school quality. Other studies document the organizational, instructional, and social adjustments that schools experience as voucher participation expands.[1] Together, this research suggests that voucher programs can meaningfully alter the environments of participating private schools. Yet we know little about whether these adjustments affect the educational outcomes of incumbent students already enrolled in private schools, nor how students themselves experience these changes. Understanding these consequences is important for evaluating the full effects of voucher programs, particularly during periods of implementation and expansion when schools may be adapting to new student populations.

One mechanism through which voucher programs could impact the educational outcomes of incumbent private school students is through the introduction of new peer groups (Epple and Romano, 1998). Students using a voucher to transition from a public to private school (from hereon referred to as voucher students) may face disruption effects, which could divert school resources away from incumbent students. Alternatively, if voucher students are academically motivated or come from families who are highly engaged, there may be a positive effect on incumbent students. More broadly, the influx of voucher students may alter classroom composition and require schools to adapt instructional practices to serve increasingly heterogeneous student populations (Carrell et al., 2018; Balestra et al., 2022; Hill et al., 2023). Understanding whether and to what extent the

[1] Notable examples include studies that examine the effects of voucher policies on private school enrollment, finances, school inputs, and quality (Rinz, 2015; Hungerman and Rinz, 2016; Adger and Felegi, 2025) as well as studies focused on the early voucher implementation experiences of private schools (Austin, 2018; Berends et al., 2023).

introduction of voucher using students impacts incumbent private school students is essential for understanding the overall effects of private school voucher programs.

The challenge with empirically examining the impact of voucher using students on the educational performance of incumbent private school students is both the lack of data on private school students and methodological limitations due to non-random sorting. Our study seeks to overcome these issues by leveraging a rich set of student-level longitudinal data, which includes demographic, achievement, and attendance information for all students in voucher-participating private schools, including all incumbent students. To deal with the selection issue, we make use of the cross-grade, within-school variation in the share of students using a voucher and show that this variation is uncorrelated with observed student characteristics including predicted test scores.

Our context centers on the Indiana Choice Scholarship Program (ICSP), one of the largest private school voucher programs in the United States, which was enacted in 2011 and first expanded in 2013. We begin with student-level testing data that covers all students in the state between the 2011-2012 and 2017-2018 academic years (AY). We use these data to construct individual-level measures of standardized test scores for non-voucher private school students and school-grade-cohort measures of the share of voucher students.[2] We define a voucher student as one that uses a voucher to attend the participating private school after having attended a public school for at least one academic year. Using this dataset, we estimate the effects of voucher students on incumbent students, using a fixed effects model that relies on the plausibly exogenous variation in the proportion of voucher students in a school by grade cohort.

We complement the administrative data analysis with an analysis of two waves of interviews conducted in nine voucher-participating private schools during the same policy period (first wave, 2012-13; second wave, 2017-19). We use these data to provide direct evidence on classroom conditions and instructional processes that are difficult to observe in administrative records. Specifically, we analyze teacher and incumbent non-voucher student interview transcripts using a structured deductive coding framework to document how educators and students describe classroom composition, instructional practices, and peer dynamics as voucher enrollment increases. This qualitative evidence provides contemporaneous insight into classroom processes and experiences that may plausibly explain any observed of voucher students on incumbent student achievement. In doing so, it helps contextualize how schools and classrooms adjust to expanding voucher participation and the ways in which these adjustments may affect students who were already enrolled in private schools.

We focus on the initial enactment and first expansion of the ICSP, which was a period during which private schools experienced substantial inflows of new students who were previously enrolled

[2]We focus our analysis on incumbent private school students in order to avoid issues of reflection Manski (1993).

in schools outside the private sector. Although Indiana further expanded its voucher program to (near-)universal student eligibility beginning in 2021, this earlier period provides a useful setting for studying the implementation dynamics that accompany the introduction and expansion of private school choice programs.[3] As many states have recently enacted or expanded similar choice programs, understanding how schools respond to potential initial enrollment shocks is highly relevant for contemporary policy debates.

From our fixed effects approach, we find that, on average, for each 10 percentage point increase in the proportion of voucher students in a given cohort reduces overall test scores between 2.00 to 2.43 percent of a standard deviation, math test scores between 2.39 and 2.61 percent of a standard deviation and ELA test scores between 1.31 and 1.78 percent of a standard deviation. The marginal effect of adding an additional voucher student to a cohort of 40 hovers around half a percent of a standard deviation across each of these outcomes of interest. However, the impact of voucher students on test scores varied significantly across student groups. While we find no evidence of heterogeneous results across gender, there are significant differences in outcomes based on whether incumbent private school students identify as non-White and their prior test scores. Students that identify as non-White see a decline of around 7 percent of a standard deviation in overall test scores for each 10 percentage point increase in the proportion of voucher students, while bottom and top tercile students see a decline between 3 and 4 percent of a standard deviation.

We further examine the impact of voucher students on incumbent private school students by exploring heterogeneity by program timing. We may expect heterogeneous results across the policy time horizon if schools mature in how they work with, integrate, and support new students, or if the composition of voucher students changes as the program expands. When we split the sample by years before and after ICSP expansion, we find that the association between the share of voucher students and academic achievement of non-voucher students is concentrated in the first two years of the program. These results suggest that participating private schools are making adjustments as the program matures. We find some evidence that voucher students entering private schools after the expansion of ICSP differed from those entering in the first two years in demographic characteristics, suggesting that schools may be doing a better job of screening out potentially disruptive students.

To understand why an increased share of voucher students has negative impacts on incumbent private school students, we document changes in school-grade characteristics. We find that higher shares of voucher students is associated with lower shares of White students, higher shares of Black students and higher shares of students that qualify for subsidized lunch, aligning with the characteris-

[3]We also note that during the recent phase of expansion beginning in 2021, most new voucher recipients in Indiana were those already enrolled in private schools who became eligible after the income eligibility threshold was expanded.

tics of voucher students themselves. Importantly, we also find that higher shares of voucher students is associated with increases in the total number of students in a school-grade cohort, suggesting that at least part of our results are explained by increasing class sizes.

Prior work shows that the attendance and discipline of classmates can have a direct impact on student's own academic achievement (Gottfried, 2009; Lavy and Schlosser, 2011). We, therefore, also examine attendance and discipline rates of voucher students as a potential mechanism for our findings. We find that voucher students miss more days of school than existing students. Specifically, the median voucher student misses an additional two days of school compared to their non-voucher counterparts. Furthermore, using data from the public schools, we also find that as ICSP matured, students that left the public school system to use a voucher were increasingly likely to have faced disciplinary action before participating in the program. By 2017, around 16 percent of eventual voucher students had ever been suspended or expelled. Together, these results suggest that increased disruptions in the classroom are a potential mechanism behind the declines in achievement.

The findings from the teacher and incumbent student interviews provide complementary evidence on the mechanisms underlying these patterns. Across interviews, teachers most consistently described changes in classroom academic composition, particularly increased heterogeneity in student preparedness, which led to greater differentiation, reteaching, and adjustments to instructional pacing. These instructional demands were more prominent than behavioral challenges, which were mentioned, but less consistently emphasized. In contrast, students primarily described changes in peer relationships and classroom climate associated with the influx of new students, with limited evidence of perceived changes in instruction or access to teacher attention. Across both teachers and students, reports of these changes were more prevalent in the early years of the program and less frequently framed as disruptions in later periods, consistent with a pattern of initial adjustment followed by adaptation.

Our results are robust to model specification choices and sample restrictions that threaten the validity of our findings. First, we assess the credibility of our main identifying assumption by showing that greater exposure to voucher students is uncorrelated to changes in incumbent students' own predicted test score, demographic characteristics, or the likelihood of leaving a participating private school. Second, we assuage concerns of insufficient variation in our main independent variable of interest by plotting the across grade variation in voucher peer share over our entire sample, as well as by year and grade-level. Third, we address the concern that our results are driven by an additional influx of new students rather than a specific effect from greater exposure to voucher students by running a falsification test where we replace the share of voucher students with the share of non-voucher new students in a school-grade-cohort. We find limited impacts of new students

on incumbent student's academic achievement, bolstering the claim that the impacts of voucher students is not solely driven by the fact that they are new to the participating private school. Last, we show our results are quantitatively similar when we expand our outcome sample to incorporate all incumbent private school students, including those that eventually use a voucher themselves.

Our paper speaks directly to the growing economics literature on private school choice programs. Much of the existing work on private school vouchers in the United States centers on the direct effects of these policies on the educational outcomes of participating students. One strand of this literature investigates whether participants realize test score gains (Rouse, 1998; Mayer et al., 2002; Howell et al., 2002; Witte et al., 2014; Wolf et al., 2010; Abdulkadiroğlu et al., 2018; Waddington and Berends, 2018), while Chingos and Peterson (2015) and Chingos et al. (2025) examine longer-run outcomes such as high school graduation, college enrollment, and college completion. We complement this body of work by showing that voucher policies also affect the educational outcomes of students who do not participate in the program. In particular, we find that as the ICSP is implemented, incumbent private school students experience declines in academic achievement. Documenting this spillover effect on incumbent students allows for a fuller accounting of the total effects of voucher policies as they are adopted and scaled.

A large body of work examines the response of public schools to the voucher programs in Milwaukee, Ohio, and Florida (Hoxby, 2003; Figlio and Rouse, 2006; Chakrabarti, 2008, 2013; Rouse et al., 2013; Chiang, 2009; Greene and Marsh, 2009; Figlio and Hart, 2014; Figlio and Karbownik, 2016; Figlio et al., 2023). Overall, these studies investigate the introduction of school voucher programs and find modest positive effects on public school performance. Our results show that while students remaining in public schools benefit from the competitive pressure vouchers create, incumbent private school students experience comparable declines in academic achievement. This symmetry suggests the welfare effects of voucher programs must consider that the benefits to public school students may be partially offset by the costs borne by incumbent private school students.

Related work outside the voucher literature examines how changes in student composition affect incumbent students. Hill et al. (2023) show that school reassignment can generate disruption and peer effects for incumbent students through the introduction of new classmates. Comparable evidence from private school voucher programs is limited. Existing research has largely focused on how participating private schools respond to voucher adoption rather than the effects on incumbent students. This area of research is understudied in the United States because of the difficulty in obtaining consistent data. Private schools, in most states, receive minimal oversight and are not subjected to the same data reporting requirements as public schools. Therefore, much of the research in this space focuses on examining responses using data on potentially selected subsets of private

schools or on publicly available datasets with limited information on academic performance.

Rinz (2015) and Hungerman and Rinz (2016) use data from the National Center for Charitable Statistics (NCCS) and the Private School Survey (PSS) to show the impact of early voucher programs on educational inputs, enrollment and fiscal outcomes at private schools. More recently, Harris and Olivier (2026), combine the NCCS and PSS with data from the Private School Review to examine the effect of voucher programs on enrollment and tuition following the expansion of universal school voucher programs since 2020. Overall, these studies find that voucher programs increase private school revenue, through the influx of public funds and higher tuition, and increase enrollment and student-teacher ratios. The finding that voucher programs are associated with higher student-teacher ratios and higher per-pupil expenditures leads to theoretically ambiguous impacts on student achievement. Our study, therefore, directly builds on the results of these papers by examining the impact of a large-scale voucher program on the academic achievement of private school students.

Adger and Felegi (2025) also study the impact of the Indiana Choice Scholarship Program (ICSP) on private school responses by evaluating the impact of the program on school quality. The authors find suggestive evidence that average private school quality fell following the adoption of ICSP, with private schools surrounded by relatively more public schools witnessing slightly larger declines. The main contribution of our paper over this prior work is the ability to examine the mechanisms that explain why private schools saw decreases in quality following the adoption of ICSP. Our findings show that ICSP had meaningful negative impacts on the academic achievement of incumbent private school students, suggesting that peer disruption and the reallocation of instructional resources are important mechanisms behind the aggregate quality declines observed at the school level.[4]

The remainder of this paper is organized as follows. In Section II, we provide background information on the Indiana Choice Scholarship Program. Section III summarizes the data used in this paper. Section IV describes the empirical strategy, lays out the regression specifications, and provides a description of the qualitative analysis. Section V contains the main results, which include our heterogeneity analysis. Section VI includes the qualitative results and connection to the quantitative findings. Section VII offers a discussion of and implications from our findings.

[4]Furthermore, Waddington and Berends (2018) show that students participating in the Indiana Choice Scholarship Program saw declines in their test scores following the switch from the public to private sector compared to matched remaining public school students. Our results show that the declines in average private school quality reported in Adger and Felegi (2025) were not solely driven by these voucher participants.

II The Indiana Choice Scholarship Program

Enacted prior to the start of the 2011-12 school year, the ICSP began as a modest, means-tested voucher program. During the first two years of the program, most first-time participants were required to have attended a public school before receiving a voucher, meaning nearly all early voucher recipients transferred from a public to a private school. Although participation caps existed initially, they were never reached, with 3,911 participating in the first year (2011-12) and 9,139 in the second year (2012-13) (Indiana Department of Education, 2016). Beginning in 2013-14, the program expanded substantially by eliminating participation caps and creating additional eligibility pathways. These pathways included siblings of participating students, kindergarten entrants, students with special learning needs, and students assigned to failing public schools (Indiana Department of Education, 2016). By 2018-19, 36,290 students were using vouchers statewide, nearly 60 percent of whom never attended a public school because of these expanded eligibility pathways (Indiana Department of Education, 2019).

Despite these expansions, the ICSP remained a means-tested program from its inception through 2020-21, with voucher eligibility determined by household income and family size. Students with household incomes at or below the federal reduced-price lunch threshold were eligible for vouchers worth up to 90 percent of the state per-pupil funding allocation in their district of residence, while those with incomes up to 150 percent of that threshold were eligible for vouchers worth up to 50 percent of the allocation.[5] In 2018-19, more than 60 percent of all Indiana families were eligible for at least a 50 percent voucher, making the ICSP one of the nation's broadest means-tested private school choice programs even before its later transition to near-universal eligibility (Waddington, 2020). This broad eligibility makes Indiana a particularly informative setting for studying how private schools adjust to the implementation and early expansion of a large-scale voucher program.

Following the COVID-19 pandemic, Indiana joined a growing number of states that expanded private school choice programs with universal or near-universal eligibility by raising income thresholds and eliminating additional eligibility requirements (Lieberman et al., 2024). In Indiana, household income eligibility increased substantially in 2021-22 (300 percent of the reduced-price lunch threshold) and again in 2022-23 (400 percent), eligibility pathways were eliminated in 2023-24, and universal eligibility will take effect beginning in the 2026-27 school year. Further, all participating

[5]The voucher amount distributed to the private school was the smaller of either the school tuition and fees or the qualified voucher amount based on per-pupil funding. During the 2018-19 school year, the average voucher amount for students in grades 1-8 was $4,965 for students qualifying for the 90% voucher and $2,923 for those receiving the 50% voucher. About 40 percent of all students received a voucher that fully covered tuition and fees at their participating private school (Indiana Department of Education, 2019).

students since 2021-22 have received 90 percent vouchers. As a result, the ICSP has become one of the largest private school choice programs in the United States, serving 76,067 students in 373 private schools as of the 2024-25 school year, with approximately 98 percent of families with school-age children eligible to participate statewide (EdChoice, 2026; Indiana Department of Education, 2025).

Our analysis focuses on the first seven years of the ICSP (2011-12 through 2017-18), which spans both the initial implementation of the program and its first major expansion. We have chosen this window for multiple reasons. First, the early years of the program were characterized by substantial student mobility from public to private schools, creating the greatest potential for disruption to existing private school environments. In later years, especially after the most recent expansions, the vast majority of new voucher recipients were already enrolled in private schools (Indiana Department of Education, 2025). Second, Indiana changed its annual assessments starting in the 2018-19 school year and did not conduct any statewide assessments in the 2019-20 school year due to the COVID-19 pandemic. Third, we have complementary qualitative data from a sample of schools at two time points early in the program (2012-13 and 2017-18 school years) that can help to contextualize these findings.

Even though our analysis focuses on Indiana's program during a period of means-tested eligibility, the implementation dynamics we study remain highly relevant to contemporary school choice policy. The central feature of the early years of the ICSP was the rapid entry of new students into participating private schools, many of whom would not otherwise have enrolled in the absence of a public subsidy. While today's universal and near-universal programs vary in their eligibility rules across states, many continue to prioritize students from lower-income families or otherwise encourage participation among students who would not have enrolled in private schools without a public subsidy. As a result, these programs likely also increase student mobility into participating private schools during periods of implementation and expansion. Consequently, the mechanisms we examine are relevant not only to Indiana's early experience in the earlier years, but more broadly to understanding how private schools adjust during the implementation and expansion of statewide private school choice programs.

III Data

III.A Indiana Department of Education Data

The quantitative data for this project come from the Indiana Department of Education's (IDOE) student longitudinal data system. We accessed these data through a longstanding data sharing agreement between the IDOE and the Center for Research on Educational Opportunity (CREO) at

the University of Notre Dame. The IDOE database contains student-level data with information on the test scores, voucher take-up, and demographics of all students enrolled in a public, private, or charter schools in Indiana. The data span from the 2005-06 through present. However, our study sample only covers the 2011-12 through 2017-18 years to coincide with ICSP's initial implementation and first expansion but prior to the statewide test changing in 2018-19.[6]

We focus on students in voucher-participating private schools that serve any students in grades 3-8. Standardized testing takes place annually in these grades in the subjects of mathematics and English/language arts. In our analyses, we standardized test scores are within year and grade, so that all estimates can be interpreted as standard deviations. Participation in annual testing and data reporting is required by voucher-participating private schools for not only voucher students, but also all non-voucher students in the school as well. These requirements are a contingent part of accepting vouchers and also required in order to remain accredited by the state (Indiana Code §20-32-5-17). Not only are the IDOE data unique because of the robust, comparable data in both public and private schools, but many private schools participated in annual statewide testing prior to the enactment of the voucher program (including most of the state's Catholic schools) for the purposes of accreditation, receiving a state accountability letter grade for marketing purposes, and participation in statewide athletics competitions. As such, for many private schools in our sample, statewide testing was not a concurrent shock as part of participating in the voucher program.

Demographic information in the IDOE data varies depending on whether a student in a private school uses a voucher. For all students, we have information on race, age, date of birth, free or reduced-price lunch status, Section 504 status, zoned school district, and standardized testing accommodations. For non-voucher private school students, we have information on whether they would qualify for a 90% voucher as it is the same cutoff for free or reduced-price lunch eligibility, which is recorded for each student. For students using a voucher, we also have information on their home address, the tuition they are charged, their pathway for qualifying, and the amount they receive. Our final dataset for the quantitative analysis consists of individual-level information on test scores and demographics for students in the third through eighth grades that were previously enrolled in a private school and had non-missing test score and demographic data.

Table 1 reports the summary statistics of incumbent private school students and voucher students. Only non-voucher private school students are in the outcome sample. Compared to students that use a voucher, non-voucher private school students are predominantly White (87.9% versus 47%), are much less likely to qualify for subsidized lunch (3.4% versus 64.8%), and use testing accommodations

[6]While our sample period relies on the academic years between 2011-2012 and 2017-2018, we do incorporate an additional year of data (the 2010-2011) to construct pre-policy school-level characteristics.

at lower rates (5.2% versus 12.1%). Furthermore, incumbent private school students outperform voucher students on standardized tests. Incumbent private school students, on average, score half a standard deviation above all students in the state, while voucher students perform around 30 percent of a standard deviation below the average in Math and 14 percent of a standard deviation below the average in English Language Arts. The differences between the two groups of students supports the idea that the ICSP introduced new groups of students into participating schools, suggesting a potential role for peer effects.

III.B Interview Sample and Data

We also used interview data from two waves of semi-structured interviews with principals, teachers, parents, and students conducted in ICSP-participating private schools in Indiana. These data were collected by a larger project team at the University of Notre Dame initially during the 2012-13 school year (first wave)[7] with a follow up spanning the entire 2017-18 and start of 2018-19 school years (second wave). The 2012-13 year corresponds to the second year after the enactment of the voucher program, and schools were adjusting to increasing enrollments of new students. The 2017-18 and 2018-19 years were well after the program's first expansion and when voucher enrollments had reached a steady state of between 35,000 and 40,000 students statewide receiving vouchers annually.

Because our primary analysis focuses on how voucher student enrollment affects incumbent private school students in grades 3–8, the most relevant interview subjects are teachers and incumbent non-voucher students in the elementary and middle testing grades. To select teachers for interviews, the Notre Dame project team stratified by grade level, ensuring that at least one teacher was selected from grades 2-3, 4-5, and 6-8. All teachers were approved by the school's principal before the project team contacted them directly to inquire about participating. If a teacher initially declined to participate, a different teacher was randomly chosen within the same stratum. To select students for participation, the project team relied on volunteers, sending a flier home with students describing the study and seeking the approval of parents to participate. Interested parents and students then directly contacted the project team to schedule an interview. The project team provided teachers and families with a nominal gift card as a remuneration for their participation.

The Notre Dame project team completed interviews in nine schools serving students in grades 3-8. Eight of the nine schools were Catholic schools and the other school was Christian-affiliated. The vast majority of voucher-participating schools in the state are religiously affiliated. The schools were spread across geographic regions, including Fort Wayne, Indianapolis, and South Bend[8]. In

[7] One of this study's authors was involved in the recruitment, protocol design, interviewing, transcription, and early analysis of this first wave of interviews in 2012-13, yielding an important contextual understanding of these data.

[8] School participation was voluntary and the interview sample was not intended to be statistically representative.

these nine schools, the project team conducted interviews with 33 teachers in the first wave. In five of these schools, they also conducted interviews with 10 non-voucher students in the first wave. In the second wave of data, the same eight Catholic schools participated in teacher interviews (25 teachers total, 10 of whom were the same as in the first wave). In six of these schools, the project team also conducted interviews with 12 non-voucher students. The Christian-affiliated school declined to participate in the second wave.

The content of the semi-structured interviews focused on the implementation, early impacts, and later steady-state of the ICSP after adding additional eligibility pathways beyond prior public school enrollment. Most relevant to this study, the interviewers asked teachers in both waves open-ended questions about instructional and curricular changes, changes in classroom dynamics, and impacts on existing students. The interviewers asked non-voucher students open-ended questions about peer interactions, integration of new students, and changes in classroom dynamics. See Appendix Section B1 for a full description of the sample and coding procedure.

IV Research Design and Analysis

IV.A Empirical Strategy

To estimate the effects of voucher students we use a fixed effects model that relies on plausibly exogenous variation in the proportion of students participating in ICSP in a school-grade-cohort. Measuring such effects has proven difficult because of both the reflection and selection problems. The reflection problem exists because incumbent and voucher student outcomes occur simultaneously, making it difficult to distinguish between the effect of a group on an individual and the effect of an individual on a group (Manski, 1993). To overcome this problem, we rely on variation in the proportion of one's school-grade-cohort who use a voucher to attend the private school. Thus, we assume that a child's peers do not cause that child to use a voucher. This assumption is met as voucher eligibility is set by a child's past education and their family's income which cannot be changed by the presence of students that use a voucher.

The selection problem arises because students self-select into schools and peer groups that are similar to them (Hoxby, 2000). In the absence of being able to randomize students into peer groups, the main approach to overcome selection has been to exploit the natural variation in cohort composition across time within a given school (Bifulco et al., 2011; Carrell et al., 2018; Balestra et al.,

However, the interview schools are broadly similar to the population of religious private schools in Indiana along observable demographic characteristics, though Catholic schools were more likely to respond and agree to participate. The project team also considered the level and share of voucher enrollment when recruiting schools.

2022; Hill et al., 2023; Figlio et al., 2024). We follow this approach and argue that while there is selection into schools, there is year-to-year variation in the proportion of students using a voucher across cohorts within the same school because of legislated changes in the requirements for the Indiana Choice Scholarship Program and natural variation in cohort composition. It is precisely this variation that we exploit in order to identify the impact of voucher students.

Specifically, we estimate the following model:

$$Y_{igst} = \alpha + \beta_1 \frac{\sum_{k \neq i}^{K} Voucher_{kgst}}{n_{gst} - 1} + X_{igst}^{'}\gamma + \lambda_{gs} + \sigma_{gt} + \epsilon_{isgt} \tag{1}$$

where i, g, s and t represent the individual, grade, school and academic year, respectively. In our preferred sample, individuals included in the regression (denoted by i) are students that never use a voucher to attend a participating private school. [9] Y_{igst} represents the outcome variables of interest test overall, Math, and English Language Arts standardized test scores. Overall test scores are calculated by taking the average of the standardized Math and English Language Arts (ELA) for each student in each grade. $\frac{\sum_{k}^{K} Voucher_{kgst}}{n_{gst}-1}$ is the proportion of students in a school-grade-cohort that use a voucher to attend the private school.[10]

We measure exposure as the share of students using a voucher that qualified because they had previously attended a public school. We denote these students as k in the regression. We focus on this group since voucher students that qualified via another track were likely already at the private school and it is unlikely that receiving a voucher (i.e., having the state pay for schooling) would impact non-voucher students through any discernible mechanism. X_{igst} is a vector of individual i's characteristics, including race, gender, subsidized lunch, and accommodation status. λ_{gs} and σ_{gt} are school-grade and grade-year fixed effects, respectively. In our main specification, we account for unobserved time-variant school characteristics by adding school-specific linear time trends. This approach has the advantage of controlling for unobserved factors correlated with time trends in school composition that may confound voucher student effects in schools. All standard errors allow for arbitrary correlation in errors at the school level.[11]

[9] The latter group of students are those that only qualified following the expansion of the program. Our results are not sensitive to this sample construction. In Appendix Table A2 we present the results using a sample that includes both never voucher students and those that received a voucher only after they had already attended the school (i.e., they qualified after the expansion of the voucher program).

[10] We measure voucher share at the cohort level as opposed to the classroom level due to potential nonrandom selection of students into classroom within a school and grade.

[11] Appendix Table A1 shows the robustness of our results to the inclusion of controls and weights.

IV.B Validity Checks

The key identifying assumption behind our empirical strategy is that there is no omitted variable that is time-varying and school specific, not captured by the school-specific time trends, is correlated with both the share of voucher students and standardized test scores, and not included in the individual-level controls. To assess the credibility that such factors do not exist, we conduct several balance checks to test whether the share of voucher students in a school-grade-cohort is systematically related to student background characteristics observable in the IDOE data.

We begin by testing whether greater exposure to voucher students is correlated to changes in predicted test scores. Specifically, we regress test scores on the full set of fixed effects and controls, excluding voucher student share, and use the estimated coefficients to predict a test score for each individual in the sample. This measure captures a linear combination of individual characteristics, where the weights are chosen to best predict test scores. Column (1) in Table 2 presents the results of this exercise and shows that predicted overall test score is uncorrelated with the share of voucher students, suggesting our identifying assumption holds in our context. Nevertheless, Columns (2)-(7) of Table 2 reports the results of a series of placebo tests checking whether the within-school-grade variation in the share of voucher students is associated with changes in students' gender, race, subsidized lunch status, and accommodations status. We choose these outcomes because these characteristics cannot be causally affected by one's peers but are likely correlated with the unobserved characteristics of other students selecting into the school. We find that exposure to voucher students is mostly unrelated to these outcomes, with exceptions for the the indicator for whether the incumbent private school student identifies as Hispanic. To ensure that this variable does not invalidate our estimates, we control for student characteristics in each of our specifications.

Columns (8)-(10) in Table 2 report the results of additional placebo tests that examine whether the share of voucher students is associated with incumbent students being dropped from the sample (because of missing student characteristics), switching schools, or missing test scores. Across the board, we find no evidence of an association between the voucher student share and these outcomes. However, Figlio et al. (2024) note these checks may not be sufficient to validate our empirical strategy if the relationship between these outcomes and the share of voucher students differs across student characteristics (e.g., if comparatively affluent families decide to leave when a large share of voucher students move into their school). We, therefore, further bolster the claim that our results are not due to differential selection of incumbent private school students remaining in the participating private schools by re-running the analysis in Columns (9) and (10) in Table 2 with interaction terms indicating whether a student identifies as White or qualifies for free-/reduced-price lunch. The

results are shown in Appendix Table A4.[12]

Another important challenge in our empirical approach is the possibility that there is insufficient variation in voucher student share across grades within school, which would attenuate our results. Panel A of Figure 1 displays the within school, across grade variation in the voucher student share across all observations in our sample and shows 33% of school-grade-year observations report at least 5 percent of students using a voucher and 15% report at least 10 percent. This pattern does not change if we examine observations separately by year (Panel B) or grade-level (Panel C). Panel D of Figure 1 shows the proportion of voucher students at the school-grade level after residualizing on cohort and school fixed effects, as well as school-specific linear time trends. This figure helps bolster the claim that there is sufficient variation in voucher student share to identify the effects we estimate.[13]

IV.C Qualitative Analysis

The qualitative component of this study used interviews with private school teachers and incumbent non-voucher students to provide contemporaneous evidence on how participating private schools experienced the introduction and expansion of Indiana's voucher program. Our analysis is a secondary deductive analysis of these interviews using a coding framework developed for the current study. Specifically, we used these interviews to assess whether teachers and students described changes in classroom conditions—including domains such as behavior and classroom management, increased need for differentiated instruction, and shifts in instructional time—that are consistent with the potential mechanisms underlying the quantitative findings. Although the interviews were conducted prior to the development of the quantitative analysis, they coincided with the same period examined in our quantitative analysis. The interviews are not used for identification or causal estimation; rather, they provide contextual evidence that helps interpret the quantitative results.

Following established approaches for integrating qualitative evidence within mixed methods research (Creswell and Plano Clark, 2018), we applied a deductive coding framework focused on identifying references to potential mechanisms underlying the observed achievement effects. The

[12] In Appendix Table A5 we present regression results comparing students that leave participating private schools (at non-transition grade levels) to those that remain in the same school-grade-yer. Contrary to the selection issue raised in Figlio et al. (2024), we find that switchers are disproportionately non-White, eligible for free/reduced-price lunch, have higher rates of special education and accommodation status, and have lower test scores than students who remain enrolled.

[13] Alternatively, in Appendix Figure A1 we plot the cross-grade, within-school range in voucher student share in each of the schools in our sample. We continue to find evidence that there is sufficient variation in our independent variable of interest.

qualitative coding was led by one author, who developed and applied the coding scheme through an iterative process, with ongoing discussion and feedback from the other author. We provide an overview of the qualitative analysis below, with additional detail in Appendix Section B1.

To begin, we initially specified the substantive domains from prior qualitative and mixed-methods research on Indiana's voucher program using these same interview data (Austin, 2018; Berends et al., 2023) to freshly code a pilot sample of teacher and student interviews. We analyzed the teacher and student interviews separately. We first assessed the validity of these initial codes by comparing them to an earlier coding of the same interviews conducted under the original coding framework developed by the Notre Dame project team. We then used these comparisons, along with researcher review of an AI-assisted deductive coding application of these same codes to the pilot group of interviews, to further refine code definitions before applying coding rules to the full set of interview transcripts. We then moved to coding the entire set of teacher and student interviews through an iterative and researcher-guided process with the support of an AI tool to apply the final deductive coding rules. In this process, we specified inclusion and exclusion criteria, including a thorough review of each coded turn of text, which led to multiple refinements and re-analyses to arrive at our final coded dataset. We treated the complete interviewee turn of speech as the unit of coding, excluding the interviewer prompts. Again, we analyzed the teacher and student interviews separately. Though we used an AI tool to aid our analytical process, we maintained full responsibility for all coding decisions, validation, and interpretation.

For teachers, the coding domains included instructional attention (e.g., differentiation, allocation of time), classroom composition (e.g., student heterogeneity, new students), behavioral management, instructional practice and pacing, and teacher workload or strain. For non-voucher students, the coding domains included academic experience (e.g., learning conditions, rigor), attention from the teacher, classroom composition and climate, social integration (integration difficulty, peer relationships), and voucher awareness.

For both teacher and student interviews, we used a dual-layer deductive coding framework. The first layer of coding captured statements reflecting change over time (before vs. after voucher implementation), comparisons across students (voucher vs. non-voucher or new vs. incumbent students), or processes linking changes in classroom conditions to instruction or learning within these domains. The second layer of coding captured the presence of domains regardless of explicit change, allowing for comparisons in prevalence across waves. The first layer contained an intentionally sparse application of codes to prioritize plausible pathways related to our quantitative findings, while the general layer captured broader classroom conditions.

We summarized the findings descriptively by reporting the share of interviews in which each

coding domain appears. For the teacher interviews, we also summarize the share of total turns referencing the specific mechanisms, as these interviews were much longer and more detailed than the student interviews (See Appendix Tables B4 and B5). We interpreted code frequencies as descriptive indicators of the prevalence of reported mechanisms rather than statistical tests. We have also included representative quotations to illustrate how interview content supports–or in some cases contrasts with–the patterns observed in the quantitative estimates. Although our summary of the interviews is not generalizable in a statistical sense, the qualitative findings offer structured, first-hand perspectives that contextualize and help explain the potential mechanisms behind the effects observed in the quantitative results.

V Results on Incumbent Private School Students

We begin by describing the estimated effects of the Indiana Choice Scholarship Program on existing private school students. The results of our main analysis are reported in Table 3. Each cell of the table represents the coefficient on the proportion of voucher students for separate regressions. Each column shows the results for an individual outcome of interest. Columns (1) and (2) present the results for overall standardized test scores, Columns (3) and (4) present the results for standardized Math scores, and Columns (4) and (6) present the results for standardized English Language Arts scores. Even columns report our specifications that include school-specific linear time trends.

We find that higher shares of voucher students in a school-grade cohort is negatively associated with academic performance of existing private school students. Moving from a cohort with no voucher students to a cohort entirely composed of voucher students reduces overall test scores between 20 to 24 percent of a standard deviation, math test scores between 24 and 26 percent of a standard deviation and ELA test scores between 13 and 18 percent of a standard deviation. However, under our preferred specification the results on ELA test scores are statistically insignificant. The marginal effect of adding an additional voucher student to a cohort of 40 hovers around half a percent of a standard deviation across each of these outcomes of interest. In our sample, the average number of voucher students in a school-grade-cohort is just over 1 and the average size of the cohort is just under 40, aligning these marginal estimates with the rollout of the program in our context.

We can begin to benchmark the magnitude of our results by comparing the size of our estimates to the effects that are obtained in studies that examine the influence of potentially disruptive children. Balestra et al. (2022) study the effect of disruptive peers on educational achievement in Switzerland, using variation in the share of children requiring special needs in school-grade-class cohorts. The authors conclude that adding one special needs student to a classroom of twenty causes

the achievement of non-special needs students to fall by 2.5% of a standard deviation. They estimate the peer effect parameter at the school-grade-classroom level, which complicates a direct comparison with our main estimate. However, when the authors re-estimate the their main analysis using the same aggregation level, they find an effect of just under 1% of a standard deviation, when adding one special needs student to a class of 20. This effect is similar in magnitude to those reported in Table 3. Our results are also similar in magnitude to those reported in Carrell et al. (2018), which studies the effect of disruptive peers on test scores (and long-run education and labor market outcomes), using variation in the share of children exposed to peers that faced domestic violence. They estimate that adding one disruptive student to a school-grade cohort of 25 reduces achievement by 0.014 standard deviations for students in grades 3 to 5.

We can also benchmark the magnitude of our estimate against related work examining the effects of voucher programs on students that participate in the program (voucher students) and those that remain in the public school. Estimates of the direct effect on participating students vary considerably in magnitude and sign. Abdulkadiroğlu et al. (2018) evaluated the Louisiana Scholarship Program using randomized lotteries and found that participation lowered math scores by 40 percent of a standard deviation, while other evaluations have found effect sizes ranging between -0.13 and 0.08 standard deviations (Rouse and Barrow, 2009). By contrast, the literature on public school students exposed to private school competition has generally found neutral to positive effects as programs mature ranging from 0.02-0.15 standard deviations (Epple et al., 2017). Our main result that a 10 percentage point increase in voucher student share reduces incumbent student achievement by 2 percent of a standard deviation sits within the range of effects on other groups of students, suggesting that the spillover effect we document, while modest in absolute terms, is economically meaningful and comparable in magnitude to the effects documented for students the literature has more traditionally focused on.

Several papers have also evaluated the impact of the Indiana Choice Scholarship Program on both participating voucher students and students remaining in the public schools. Papers examining ICSP's effects on traditional public school students find a range of estimates from -0.05 to 0.019 standard deviations (Canbolat, 2021; Egalite and Catt, 2025; Adger and Felegi, 2025), with most estimates clustering toward modestly positive effects. Waddington and Berends (2018) find that voucher students themselves experienced a considerably larger 0.15 SD decline in math achievement during their first year in a private school, with no significant effect on English Language Arts. Our estimated effect on incumbent private school students' math scores is roughly an order of magnitude larger than the estimates reported for public school students, and comparable in magnitude to the direct effect on voucher students themselves, suggesting that policymakers are missing a meaningful

piece of the picture when they do not consider the spillover effects experienced by incumbent private school students.

A potential concern is that our estimates reflect general adjustment costs associated with an influx of new students into a private school, rather than the specific effect of greater exposure to voucher students. To assess this possibility, we conduct a falsification test in which we re-estimate Equation 1 replacing the voucher exposure measure with the share of students who are new to the school within a given school–grade cohort. If our main results were driven by general entry effects rather than an impact from voucher students, we would expect to observe similar patterns in this falsification exercise. Instead, Appendix Table A3 shows that we find no statistically significant effect of the share of new students in a school-grade cohort on our outcome variables of interest. These results provide support for the claim that our results are not solely driven by the fact that voucher students are new to the private school.

V.A Mechanisms

Given the declines we find in test scores, we next examine changes in cohort characteristics and class sizes. We begin by documenting the link between higher shares of voucher students and the demographic composition of school-grade-cohorts. Table 4 reports the results of Equation 1 where the outcomes of interest of school-grade-year measures of demographic variables (Percent Female, Percent White, Percent Hispanic, etc.).[14] Table 4 shows that higher shares of voucher students is associated with a statistically significant decrease in the percent of White students, and statistically significant increases in the percent of Black, Hispanic and students qualifying for reduced price lunch. These results are in-line with the characteristics of voucher students as shown in Table 1.

Column (8) of Table 4 also shows that, while statistically insignificant, higher shares of voucher students is associated with larger cohort sizes, suggesting that the declines in the academic achievement of existing private school students may be driven in part by increasing class sizes. Krueger (1999) finds that a reduction of 7 students (from a class size of 22 - a 32% decrease) increases test scores by 20 percent of a standard deviation. Using this estimate, we can perform a back-of-the-envelope calculation to understand how much of our findings can be explained by increased class sizes alone. We find an increase of $\sim$ 4% in class sizes (1.5 off a base mean of 40). The Krueger (1999) estimates suggests an increase of 4% in class sizes would reduced test score by around 2.5% of a standard deviation; about one-eighth of the impact we estimate in our preferred specification in Table 3. While significant, these results suggest that the impact of voucher students on incumbent

[14] Importantly, these variables are calculated including both incumbent private school students and the voucher students themselves.

private school students are not only the result of the downstream effects of increased class sizes. Moreover, the findings on class sizes contrasts the interview data, in that no teachers reported any concerns or issues with class sizes having increased to the degree that it is impacting their instruction.

Prior work has shown that both attendance and discipline influence the academic outcomes of students (Gottfried, 2011; Lavy and Schlosser, 2011). If voucher students have higher rates of absenteeism than their incumbent private school peers, teachers may be forced to split their time between students who have and have not missed the previous lessons, reducing the academic achievement of incumbent private school students. Similarly, if voucher students have high discipline rates, this could lead to increased disruptions in class that reduces achievement. We, therefore, use attendance and discipline data in the IDOE-CREO database to evaluate differences in attendance and suspensions/expulsions by voucher status. Unfortunately, information on suspensions and expulsions is only available for public school students. Therefore, the comparisons we make involving discipline are across public school students in third through eighth grade that either go on to use a voucher or never use a voucher.

Panel A of Figure 2 plots the distribution of school days attended by voucher status. We find that voucher students miss more days of school than incumbent private school students, suggesting that our results could be driven by the differences in attendance. Panel B of Figure 2 plots the share of students between third through eighth grade that are suspended or expelled in a given year by eventual voucher status. Specifically, in 2007 only 5 percent of students that eventually left a public school to use a voucher were suspended or expelled. That number increased to 16 percent in 2017. We see an opposite pattern for remaining students. In 2007, 12 percent of remaining public school students in the third through eighth grade were suspended or expelled and that number decreased to around nine percent in 2017. The increasing rate of suspensions/expulsions of eventual voucher students suggests that they may serve as increasingly disruptive peers, potentially explaining the declines in test scores of non-voucher private school students.

V.B Heterogeneity by Student Characteristics

We have found consistent evidence that greater exposure to voucher students is associated with declines in test scores. However, these average estimates across all students could differ across various subgroups. Therefore, we disaggregated the results by the following student characteristics: gender, race and prior test performance. We calculated these estimates by introducing interactions of the student characteristics with the $\frac{\sum_k^K Voucher_{kgst}}{n_{gst}-1}$ variable in Equation 1. Table 5 displays the results of our heterogeneity analysis by student subgroup for our three outcome measures of interest: overall test score, math test score, and ELA test score, respectively.

Panel A reports the differences in outcomes for male and female students. The interaction term in Panel A isolates the differential impact of greater exposure to voucher students for those that identify as female compared to their male counterparts. We may expect differences in the impact of voucher students across gender because prior work shows that boys' academic outcomes can be more sensitive to disruptive classroom peers than girls (Lavy et al., 2012; Carrell et al., 2018; Bietenbeck, 2020), which could generate differential effects if voucher students are more likely to exhibit disruptive behavior. Despite this theoretical reason to expect heterogeneous effects, the estimate on the interaction term with identifies as Female is statistically insignificant. This result implies that Female students see similar reductions in test score performance as their male counterparts.

In Panel B, we examine whether the impact of voucher students varies by race. The interaction term in Panel B isolates the differential impact of greater exposure to voucher students for non-White students compared to their White counterparts. Non-White incumbent students may be more adversely affected by the introduction of voucher students since voucher students are disproportionately non-White (see Table 1), so their arrival changes the racial composition of classrooms (as shown in Table 4), and prior work suggests that students tend to be more sensitive to peers of the same race (Hoxby, 2000). Consistent with this hypothesis, we find that the estimate on the interaction term with identifying as non-White is negative and statistically significant. These findings suggest that the decline in academic performance is larger (more negative) for non-White students.

Panel C reports the effect of voucher students on test scores by previous academic performance.[15] Specifically, we include two interaction terms. The first identifies the interaction between exposure to voucher students and whether a student's prior year test score was in the bottom tercile of the overall test score distribution. The second is an interaction term between voucher exposure and whether a student was in the top tercile of prior year academic performance. There are reasons to expect effects at both ends of the distribution. Low-achieving incumbent students may be particularly vulnerable if voucher students draw school resources that would have otherwise been directed toward them. On the other hand, high-achieving students may be adversely affected if the introduction of voucher students slows instructional pace or reduces the complexity of classroom material as teachers adjust to a wider range of student preparation levels.

Indeed, we find that incumbent students at both the top and bottom of the prior achievement distribution experience larger declines in math test scores relative to those in the middle.[16] For ELA, the negative effects are concentrated among low-achieving incumbent students. The differen-

[15] Because we only have previous test performance for those in the fourth grade and above, this analysis necessarily drops students in the third grade.

[16] The math score results hold if we instead consider whether a student is the top or bottom tercile of the test score distribution at their own school. See Appendix Table A6.

tial results across subjects may stem from the greater school-dependence of math relative to ELA, as high-achieving students may be better positioned to supplement ELA instruction through independent reading and home learning.

In Appendix Figure A2 we explore whether the impacts of voucher students vary by the grade level of incumbent private school students. One may expect differential effects if voucher students made the switch into a participating private school in a natural transition year (i.e., sixth grade in most Indiana contexts) versus a potentially more disruptive off-transition grade, or if the classroom environment itself differs substantially across grades. Instead, we find that our point estimates are most negative across the fifth through seventh grades (albeit each is statistically insignificant at conventional levels). One possible explanation is that grades five through seven often mark a shift from single self-contained classrooms to departmentalized instruction across subjects, broadening incumbent students' exposure beyond a single classroom roster and potentially making composition changes more salient (Wellington et al., 2024; Backes et al., 2025).

We also examine whether the impacts of voucher students differs across baseline school characteristics. Schools with a below median, baseline non-White or free/reduced-price lunch share of students may experience larger disruptions from the demographic changes brought by voucher students. On the other hand, schools with an above median, baseline enrollment or school value-added may be better equipped to integrate voucher students through more experienced teachers or stronger administrative capacity. To answer this question, we disaggregated our baseline results by the following pre-policy school characteristics: enrollment of non-White students, FRPL students, total enrollment and school value-added. We calculated these estimates by splitting the sample by whether a school fell above/below the median value in the 2010-2011 academic year. Figure 4 displays the results of this heterogeneity analysis. Each panel of the figure, for a given pre-policy characteristic, displays the coefficient on the share of voucher students separately for schools above and below the median baseline value. Contrary to our hypotheses, we find that participating private schools with initially above-median shares of non-White and FRPL students saw larger declines in test scores, with no differences based on a school's baseline enrollment or quality.

V.C Heterogeneity by Voucher Student Characteristics

We next examine the heterogeneity effects by the prior academic performance of the voucher students. Results are shown in Table 6, where we show our results across three panels: impacts on overall test scores are reported in Panel A, math test scores in Panel B and English language arts test scores in Panel C. The first column for each outcome replicates our main estimates from the even columns reported in Table 3. Results from column 2 indicate that the negative effects of voucher

student exposure are substantially larger when voucher students' prior-year test scores fall in the bottom tercile of the achievement distribution. These results are roughly double the magnitude of our main estimates, and are statistically significant across all three outcomes. Conversely, for voucher students in the top tercile of the prior-year achievement distribution (column 3), the point estimates are positive and not statistically significant at conventional levels across all three outcome measures. Together, these results suggest that the negative impacts of voucher students may stem from a widening range of academic preparation in the classroom or from lower-performing voucher students requiring additional resources that come at the expend of incumbent students.

V.D Heterogeneity by Program Timing

The Indiana Choice Scholarship Program was adopted and expanded in two separate academic years (2011-2012 and 2013-2014, respectively) during our sample period. One might then wonder if the different program structures had differential effects on existing private school students. We might expect differential results by program timing if schools adjust as the program matures - either by changing school resources or the set of voucher students they admit into the school.

We answer this question by splitting the sample by years before and after the expansion. Table 7 reports the results by program timing for each outcome variable of interest. Odd columns report the coefficients when estimated using data from 2012 and 2013, while even columns report the coefficients when estimated using data from 2014 onward. We find that the association between the share of voucher students and academic performance of existing private school students are concentrated in the first two years of ICSP. All coefficients in the years post-expansion are statistically indistinguishable from zero. Furthermore, even if we took the estimated coefficients seriously, they are between a fifth to a half the size of estimates in the pre-expansion time period.

To understand why we see dissipating effects of voucher students as the program expands, we begin by examining differences in voucher student characteristics across the two time periods. Figure 3 plots the test score distribution of voucher students that entered participating private schools before versus after the expansion of the Indiana Choice Scholarship Program. Specifically, the figure plots the distribution of test scores for these students in the year before they use the voucher. The two distributions are statistically indistinguishable from one another (the p-value for the Kolmogorov-Smirnov equality-of-distributions test is equal to 0.632), suggesting that the reduction in the impacts of voucher students is unlikely to be due to participating private schools selecting more academically prepared students as the program matures.

Beyond test scores, Table 8 reports the summary statistics of voucher students that moved from the public to private sector in the first two years of ICSP (Columns (1) and (2)) versus after the

expansion (Columns (3) and (4)). Compared to students that used the voucher between 2012 and 2013, newer cohorts were less likely to identify as White (45.8% versus 48.1%) or Black (19.4% versus 21.7%), were more likely to identify as Hispanic (26.1% versus 20.5%) , and use testing accommodations (16.4% versus 13.9% for Math and 16.7% versus 13.9% for ELA). Interestingly, these demographic characteristics are more dissimilar to the average statistics of non-voucher students shown in Table 1. We also continue to find that newer voucher students performed no better on standardized tests prior to their move to the participating private schools. Voucher students entering after the expansion of ICSP performed, on average, a statistically insignificant 1 percent of a standard deviation below the 2012 and 2013 cohorts on both math and ELA. We take these differences in student characteristics over time as evidence that participating private schools adjusted the demographic composition of voucher students after the program expanded, which may play a role in the change of their impact on incumbent students.

VI Results from Teacher and Student Interviews

VI.A Teachers

We next examined teacher interviews to assess potential mechanisms through which increased voucher enrollment may affect incumbent students. Teachers generally did not frame their responses explicitly in terms of the voucher program itself, as teachers indicated varied knowledge of which students in their classroom were receiving vouchers. Instead, they nearly universally described classroom changes in terms of "new students," differences in instruction or preparation compared to prior years, or broader shifts in classroom composition. Across interviews, teachers most consistently emphasized instructional and organizational mechanisms, particularly those related to the variation in student preparedness and ability level along with the resulting demands on instruction.

Nearly all teachers in the first wave referenced changes in classroom composition, most commonly in terms of wider variation in academic readiness. Teachers frequently described incoming students as arriving with gaps in foundational skills, particularly in math and reading. For example, one teacher noted that "*when they come in... they can be a little bit lower for math, especially basic skills*," highlighting how differences in prior preparation shape classroom instruction. Another teacher described a student in which "*he's in eighth grade, but he's working at about a fifth grade level.*" Yet another teacher stated, "*some of the students... if they don't have the basics it's really hard to get them into pre-algebra or algebra.*" These differences in academic preparation required teachers to adjust how and what they taught.

Teachers consistently described responding to these differences in student preparation through

greater differentiation, reteaching, and adjustments to instructional pacing. They described implementing strategies such as ability grouping, small-group instruction, and targeted remediation, many of which had not previously been regular parts of their instruction. These changes also required teachers to spend more time with some students or material than they had in the past. While behavioral and disciplinary challenges were also present, teachers mentioned them less consistently than differences in academic preparation and the resulting instructional changes. Teachers also mentioned increased workload and strain, though these concerns were often connected to the additional demands of adapting instruction.

Comparing across waves, teachers were less likely in the second wave to describe these classroom conditions as changes or disruptions. Teachers in the first wave more frequently described changes in classroom composition and the need to adjust instruction, pacing, and grouping. In the second wave, teachers continued to discuss these same practices, but more often described them as part of their regular approach to instruction rather than a new set of practices. This pattern is consistent with an initial period of disruption followed by adaptation as schools and teachers adjusted to a more academically diverse student population.

The general layer of coding provides additional evidence of this shift. Teachers in both waves continued to discuss differentiation and variation in student academic background even when they did not describe these conditions as new or changing. For example, teachers in the second wave often discussed ability grouping and targeted support as regular features of instruction rather than as responses to recent changes. In fact, all references to the term "ability grouping" occurred in the second wave, and appeared in interview turns from teachers in five of the eight schools. These findings suggest that changes in student preparation initially required teachers to make meaningful instructional adjustments. Over time, these practices remained, but teachers increasingly described them as part of their regular instruction rather than as responses to disruption.

VI.B Students

We next examined student interviews first to assess potential mechanisms through which new students in the school using vouchers may affect incumbent students. Similar to the teachers, incumbent students had limited knowledge of which classmates were receiving vouchers and generally described their experiences in terms of "new students" rather than "voucher students." Unlike teachers, however, students focused primarily on changes in friendships, peer relationships, and the classroom environment rather than changes in instruction.

Nearly all students within the first wave referenced changes in classroom composition or climate. For example, one student explained that "*some students get left out now, because there's more*

students in our class," directly linking increased enrollment to changes in peer interactions. Another student similarly described how "*people keep a tighter circle... they're not friends with everybody anymore*," suggesting that classroom changes are associated with more fragmented peer groups. These shifts are likely related to challenges of social integration. Several students described difficulty in forming or maintaining friendships in the presence of new students. As one student notes, "*trying to find new friends... when you change schools is the hardest thing to do*," highlighting the friction associated with changes in peers.

In contrast, we found limited evidence of direct evidence of changes in instruction as perceived by students. Although students may have described how teachers structured access to help (e.g., "*you can just go up and ask the teacher what you don't understand*"), students also provided little evidence that the arrival of new students reduced their access to teachers. Although students frequently described how they could seek help from teachers, very few reported having to compete with other students for teacher attention. Similarly, few students described changes in academic rigor or instructional pace associated with changes in their classrooms, though these may also be instructional concepts that are beyond their comprehension.

Comparing across waves, the prevalence of statements associated with the mechanisms we assessed decreased substantially. Mentions of classroom composition related to new students fell from 90% in the first wave to 33% in the second wave, and references to integration difficulty fell from 60% to 33%. Similar to the findings from the teacher interviews, although through different mechanisms, this pattern is consistent with what appeared to be an initial period of disruption followed by adaptation. It appears the impact of an increasing number of new students in the classroom became either less salient or was at least perceived by incumbent students as less disruptive over time.

The broader coding again suggested that these experiences did not disappear over time. Students in both waves continued to discuss classroom composition, peer relationships, and teacher support at high rates even when they did not describe them as changes associated with the voucher program. For example, students in the second wave described peer relationships in more settled terms, with one student noting that "*it's pretty easy [to make friends] because everybody's friendly.*" Taken together, the student interviews suggest that changes in peer relationships and classroom dynamics were more disruptive to incumbent students early in the program. By the second wave, these dynamics were less present and other adjustments were more often described as normal features of their school experience.

VI.C Connection Between Peer Effects and Interview Findings

When taken together, the quantitative and qualitative findings align most clearly around mechanisms related to changes in classroom composition and the resulting instructional demands placed on teachers. The quantitative results show that higher shares of voucher peers are associated with declines in incumbent students' achievement, particularly in the early years of the program and among students potentially most sensitive to instructional changes. The interview data provided direct evidence of how these mechanisms operate in practice. Teachers consistently described increased heterogeneity in student preparedness, which led to greater differentiation, reteaching, and adjustments to pacing. Many teachers discussed these challenges in the context of mathematics instruction, where variation in prior knowledge made it more difficult to maintain a common pace and sequence of instruction. These accounts provide a plausible explanation for the stronger effects observed in mathematics, a subject in which cumulative learning and instructional sequencing may make achievement more sensitive to changes in classroom composition and instructional practices.

At the same time, the qualitative findings both refined and complicated the mechanisms suggested by the quantitative analysis. While the administrative data point to attendance and disciplinary differences as potential sources of disruption, teachers rarely centered these engagement and behavioral challenges in their accounts, instead emphasizing academic preparedness and the need for instructional adaptation. Student interviews point in another different direction, as students reported few changes in instruction or access to teacher attention and instead focused on shifts in peer relationships and classroom climate, which could be more directly related to behavior. These differences suggest that no single mechanism fully explains the quantitative results. Rather, the negative achievement effects likely reflect a combination of channels: some role for disruption, as implied by the administrative data, alongside more pervasive changes in instructional pacing and classroom dynamics that are more consistently reflected in the interview evidence.

Finally, both the quantitative and qualitative evidence point to a common pattern of initial disruption followed by adaptation. Negative effects are concentrated in the early years of the voucher program and attenuate over time. Similarly, earlier interviews emphasize instructional disruptions associated with new students, while later interviews frame these conditions as normalized features of the classroom. This convergence strengthens the interpretation that schools adjust to the introduction of voucher students, reducing the magnitude of negative effects experienced by existing students over time while the underlying instructional and social dynamics remained similar.

VII Discussion and Conclusion

This paper examines how the initial implementation and expansion of the Indiana Choice Scholarship Program affected incumbent students in participating private schools. We find that exposure to greater shares of voucher peers within a school-grade cohort is associated with declines in incumbent students' academic performance and vary by student race/ethnicity and prior academic performance. Our estimates are modest in magnitude, but they fall in the range of other papers evaluating the spillover effects of potentially disruptive peers (Carrell et al., 2018; Balestra et al., 2022). The concentration of these effects in mathematics is also consistent with Waddington and Berends (2018), who find larger negative impacts on mathematics achievement for voucher participants in the Indiana Choice Scholarship Program. However, we also find the declines are concentrated in the initial two years of the programs suggesting that there are not persistent declines in performance over time associated with the introduction of voucher students.

We provide complementary qualitative evidence that helps explain three central quantitative findings: the decline in incumbent student achievement associated with the influx of voucher students, the larger effects in mathematics, and the concentration of these effects during the initial implementation of the program. Teachers consistently described increased heterogeneity in student academic preparedness, shifts in instructional pacing and differentiation, and changes in classroom dynamics following the introduction of voucher students, while incumbent students described a more disruptive learning environment. Together, these accounts suggest that schools experienced short-run implementation challenges as they adapted to new student populations, providing a plausible explanation for both the larger effects in mathematics and the attenuation of these overall effects over time.

These results contribute to a growing literature on private school choice by shifting attention to outcomes within participating private schools. Prior work has largely focused on the effects of vouchers on participating students or on students remaining in public schools, often highlighting mixed effects for the former and modest competitive effects for the latter. Our findings extend this literature by showing that voucher programs also generate meaningful, negative, and short-term spillover impacts on non-participating students within private schools. In this sense, our results suggest that the total effects of voucher policies must not account only for the gains or losses among participants or competitive responses in public schools, but also for the adjustment costs borne by private schools during policy implementation.

We believe our results are particularly relevant in the current policy environment, where policymakers across states have either recently enacted or expanded private school choice programs (vouchers, tax credit scholarships, and educational savings accounts) or are considering doing so.

The passage of the federal school choice tax credit through the One Big Beautiful Bill Act also has implications beyond these individual states as there may be motivation for states to opt into a federal tax credit program that might otherwise not enact a statewide choice program. Although our analysis focuses on Indiana's program during the time in which it was means-tested (it is now near-universal), the core mechanisms related to student mobility we identify are likely to apply broadly to new and expanded programs in other states, even with a (near-)universal eligibility provision. In these contexts, participating schools will still likely face an influx of new students irrespective of policy design, and thus may be susceptible to similar short-run disruptions as we described here.

In particular, states with new or expanded private school choice programs should give close attention to integrating new students in private schools who have varied academic backgrounds and prior schooling experiences. Our results also caution policymakers and school leaders against interpreting early disruptions as persistent failures or responding by restricting private school access to students perceived as more challenging to serve. Instead, our evidence suggests that schools can and do respond to these initial challenges and changes over time as new practices become embedded in routine instruction and school culture.

We also note some limitations with our study. Our estimates capture average effects on incumbent students within one state and program context and may not generalize to all private school settings or choice policy designs. While we assuage identification concerns through our empirical strategy and falsification tests, we cannot fully rule out all sources of unobserved variation. And although the qualitative evidence provides valuable insight into potential mechanisms, it is not designed to establish causal pathways or represent all participating schools. Despite these limitations, this study provides some of the first systematic evidence on how voucher programs impact incumbent students within private schools and highlights the importance of considering implementation and adjustment dynamics when evaluating the broader impacts of school choice policies.

References

Abdulkadiroğlu, A., P. A. Pathak, and C. R. Walters (2018). Free to Choose: Can School Choice Reduce Student Achievement? *American Economic Journal: Applied Economics 10*(1), 175–206.

Adger, C. and B. Felegi (2025). Supply-Side Responses in School Choice. Working paper, Available at SSRN 4603968.

Austin, M. (2018). Organizational and social costs of schools' participation in a voucher program. In *School choice at the crossroads*, pp. 32–53. Routledge.

Backes, B., J. Cowan, and D. Goldhaber (2025). Building bridges to middle school? elementary school departmentalization and academic achievement in the upper grades. CALDER Working Paper 315-0325, Center for Analysis of Longitudinal Data in Education Research (CALDER), American Institutes for Research.

Balestra, S., B. Eugster, and H. Liebert (2022). Peers with Special Needs: Effects and Policies. *Review of Economics and Statistics 104*(3), 602–618.

Berends, M., R. J. Waddingtion, and M. Austin (2023). School Vouchers in Indiana: Policy Shifts and their Implications for Economically Disadvantaged Families and Students of Color. In *The School Voucher Illusion: Exposing the Pretense of Equity*. Teachers College Press.

Bietenbeck, J. (2020). The Long-Term Impacts of Low-Achieving Childhood Peers: Evidence from Project STAR. *Journal of the European Economic Association 18*(1), 392–426.

Bifulco, R., J. M. Fletcher, and S. L. Ross (2011). The Effect of Classmate Characteristics on Post-Secondary Outcomes: Evidence from the Add Health. *American Economic Journal: Economic Policy 3*(1), 25–53.

Canbolat, Y. (2021, 7). The Long-Term Effect of Competition on Public School Achievement: Evidence from the Indiana Choice Scholarship Program. *Education Policy Analysis Archives 29*, 97.

Carrell, S. E., M. Hoekstra, and E. Kuka (2018, November). The Long-Run Effects of Disruptive Peers. *American Economic Review 108*(11), 3377–3415.

Chakrabarti, R. (2008). Can Increasing Private School Participation and Monetary Loss in a Voucher Program Affect Public School Performance? Evidence from Milwaukee. *Journal of Public Economics 92*(5-6), 1371–1393.

Chakrabarti, R. (2013). Vouchers, Public School Response, and the Role of Incentives: Evidence from Florida. *Economic Inquiry 51*(1), 500–526.

Chiang, H. (2009). How Accountability Pressure on Failing Schools Affects Student Achievement. *Journal of Public Economics 93*(9-10), 1045–1057.

Chingos, M., D. Figlio, and K. Karbownik (2025). The Effects of Ohio's EdChoice Voucher Program on College Enrollment and Graduation. Research report, Urban Institute.

Chingos, M. M. and P. E. Peterson (2015). Experimentally Estimated Impacts of School Vouchers on College Enrollment and Degree Attainment. *Journal of Public Economics 122*, 1–12.

Creswell, J. W. and V. L. Plano Clark (2018). *Designing and Conducting Mixed Methods Research (3rd ed.)*. SAGE Publications, Inc., Thousand Oaks, CA.

EdChoice (2026). The ABCs of School Choice: The Comprehensive Guide to Every Private School Choice Program in America. 2026 Edition. Technical report, EdChoice.

Egalite, A. J. and A. D. Catt (2025, 3). Effects of the Indiana Choice Scholarship Program on Public School Students' Achievement and Graduation Rates. *AERA Open 11*, 2.

Epple, D. and R. Romano (1998). Competition between Private and Public Schools, Vouchers, and Peer-Group Effects. *American Economic Review 88*(1), 33–62.

Epple, D., R. E. Romano, and M. Urquiola (2017). School Vouchers: A Survey of the Economics Literature. *Journal of Economic Literature 55*(2), 441–92.

Figlio, D., P. Giuliano, R. Marchingiglio, U. Ozek, and P. Sapienza (2024). Diversity in Schools: Immigrants and the Educational Performance of US-born Students. *Review of Economic Studies 91*(2), 972–1006.

Figlio, D. and C. Hart (2014). Competitive Effects of Means-Tested School Vouchers. *American Economic Journal: Applied Economics 6*(1), 133–56.

Figlio, D. and K. Karbownik (2016). Evaluation of Ohio's EdChoice Scholarship Program: Selection, Competition, and Performance Effects. Technical report, Thomas B. Fordham Institute.

Figlio, D. N., C. Hart, and K. Karbownik (2024). Competitive Effects of Charter Schools. Technical report, National Bureau of Economic Research.

Figlio, D. N., C. M. Hart, and K. Karbownik (2023). Effects of maturing private school choice programs on public school students. *American Economic Journal: Economic Policy 15*(4), 255–294.

Figlio, D. N. and C. E. Rouse (2006). Do Accountability and Voucher Threats Improve Low-Performing Schools? *Journal of Public Economics 90*(1-2), 239–255.

Gottfried, M. A. (2009). Excused Versus Unexcused: How Student Absences in Elementary School Affect Academic Achievement. *Educational Evaluation and Policy Analysis 31*(4), 392–415.

Gottfried, M. A. (2011). Absent Peers in Elementary Years: The Negative Classroom Effects of Unexcused Absences on Standardized Testing Outcomes. *Teachers College Record 113*(8), 1597–1632.

Greene, J. P. and R. H. Marsh (2009). The Effect of Milwaukee's Parental Choice Program on Student Achievement in Milwaukee Public Schools. SCDP Comprehensive Longitudinal Evaluation of the Milwaukee Parental Choice Program. Report# 11. Research report, School Choice Demonstration Project.

Harris, D. and G. Olivier (2026). The Effects of Universal School Vouchers on Private School Tuition and Enrollment: A National Analysis. Technical report, National Center for Research on Education Access and Choice.

Hill, D. V., R. P. Hughes, M. A. Lenard, D. D. Liebowitz, and L. C. Page (2023). New Schools and New Classmates: The Disruption and Peer Group Effects of School Reassignment. *Economics of Education Review 92*, 102316.

Howell, W. G., P. J. Wolf, D. E. Campbell, and P. E. Peterson (2002). School Vouchers and Academic Performance: Results from Three Randomized Field Trials. *Journal of Policy Analysis and Management 21*(2), 191–217.

Hoxby, C. M. (2000). Peer Effects in the Classroom: Learning from Gender and Race Variation. Working paper 7867, National Bureau of Economic Research.

Hoxby, C. M. (2003). School Choice and School Competition: Evidence from the United States. In *School Choice and School Competition: Evidence from the United States*.

Hungerman, D. M. and K. Rinz (2016). Where Does Voucher Funding Go? How Large-Scale Subsidy Programs Affect Private-School Revenue, Enrollment, and Prices. *Journal of Public Economics 136*, 62–85.

Indiana Department of Education (2016). Choice Scholarship Program Annual Report, Participation and Payment Data. https://www.in.gov/doe/files/2015-2016-Annual-Report.pdf [Accessed: 2024-11-13].

Indiana Department of Education (2019). Choice Scholarship Program Annual Report, Participation and Payment Data. https://www.in.gov/doe/files/2018-2019-Annual-Report.pdf [Accessed: 2024-11-13].

Indiana Department of Education (2025). Choice Scholarship Program Annual Report, Participation and Payment Data. https://www.in.gov/doe/files/2024-2025-Annual-Choice-Report.pdf [Accessed: 2026-07-24].

Krueger, A. B. (1999). Experimental Estimates of Education Production Functions. *The Quarterly Journal of Economics 114*(2), 497–532.

Lavy, V., M. D. Paserman, and A. Schlosser (2012). Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom. *The Economic Journal 122*(559), 208–237.

Lavy, V. and A. Schlosser (2011, April). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics 3*(2), 1–33.

Lieberman, M., L. Stanford, and V. A. Ifatusin (2024). Which States Have Private School Choice? https://www.edweek.org/policy-politics/which-states-have-private-school-choice/2024/01 [Accessed: 2026-07-24].

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies 60*(3), 531–542.

Mayer, D. P., P. E. Peterson, D. E. Myers, C. C. Tuttle, and W. G. Howell (2002). School Choice in New York City After Three Years: An Evaluation of the School Choice Scholarships Program. Final Report. Research report, Mathematica Policy.

Rinz, K. (2015). Undone by the Market? The Effects of School Vouchers on Educational Inputs. Unpublished Working Paper.

Rouse, C. E. (1998). Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program. *The Quarterly Journal of Economics 113*(2), 553–602.

Rouse, C. E. and L. Barrow (2009). School Vouchers and Student Achievement: Recent Evidence and Remaining Questions. *Annual Review of Economics 1*(1), 17–42.

Rouse, C. E., J. Hannaway, D. Goldhaber, and D. Figlio (2013). Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure. *American Economic Journal: Economic Policy 5*(2), 251–81.

Waddington, R. J. (2020). Social context of vouchers. In *Handbook of research on school choice (2nd ed.)*, pp. 202–213. Routledge.

Waddington, R. J. and M. Berends (2018). Impact of the Indiana Choice Scholarship Program: Achievement Effects for Students in Upper Elementary and Middle School. *Journal of Policy Analysis and Management 37*(4), 783–808.

Wellington, A., M. Clark, A. Burnett, S. James-Burdumy, L. Makowsky, S. Brockman, D. Dotter, M. Herrmann, and H. Chiang (2024, 7). Evaluation of Departmentalized Instruction in Elementary Schools: Exploring Implementation Experiences. NCEE Report 2024-005, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Prepared by Mathematica, Public Impact, Social Policy Research Associates, and Clowder Consulting under contract ED-IES-17-C-0064.

Witte, J. F., P. J. Wolf, J. M. Cowen, D. E. Carlson, and D. J. Fleming (2014). High-Stakes Choice: Achievement and Accountability in the Nation's Oldest Urban Voucher Program. *Educational Evaluation and Policy Analysis 36*(4), 437–456.

Wolf, P., B. Gutmann, M. Puma, B. Kisida, L. Rizzo, N. Eissa, and M. Carr (2010). Evaluation of the DC Opportunity Scholarship Program: Final Report. NCEE 2010-4018. Research report, National Center for Education Evaluation and Regional Assistance.

VIII Tables

Table 1: Descriptive Statistics

	Non-Voucher Students		Public School Voucher Students		
	Mean	SD	Mean	SD	**Difference**
Panel A: Demographic Characteristics					
White	0.879	(0.326)	0.470	(0.499)	0.409***
Black	0.018	(0.132)	0.169	(0.375)	-0.151***
Hispanic	0.048	(0.214)	0.271	(0.445)	-0.223***
Female	0.501	(0.500)	0.518	(0.500)	-0.017***
FRPL	0.034	(0.182)	0.648	(0.478)	-0.613***
Has Math Accommodation	0.052	(0.223)	0.121	(0.326)	-0.069***
Has ELA Accommodation	0.052	(0.223)	0.120	(0.325)	-0.068***
Panel B: Education Outcomes					
Standardized Test Score	0.476	(0.820)	-0.245	(0.832)	0.721***
Standardized Math Score	0.393	(0.831)	-0.313	(0.840)	0.705***
Standardized ELA Score	0.506	(0.850)	-0.136	(0.852)	0.642 ***

Notes: This table shows descriptive statistics about non-voucher and voucher students at participating private schools throughout the sample time frame. Data on enrollment and test scores come from the IDOE-CREO database. Non-voucher students are defined as incumbent private school students that never attended a public school, while public school voucher students are those that used the voucher to move from the public to private sector. Column (5) reports the difference in the means between public school voucher students and non-voucher students.

Table 2: Results on Exogenous Student Characteristics and Attrition

VARIABLES	Predicted Test Score	Female	White	Black	Hispanic	FRPL	Any Accomm.	Drop Sample	Change Schools	Missing Scores
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Exposure to Public School Voucher Students										
Fraction Students using Voucher	0.028	0.069	-0.056	-0.003	0.071**	0.000	0.030	0.018	-0.100	-0.014
	(0.091)	(0.052)	(0.044)	(0.020)	(0.032)	(0.030)	(0.024)	(0.045)	(0.073)	(0.014)
	[0.001]	[0.002]	[-0.001]	[< -0.001]	[0.002]	[0.000]	[0.001]	[< 0.001]	[-0.002]	[< -0.001]
Observations	92,549	113,618	113,618	113,618	113,618	113,618	113,618	118,889	74,574	114,332
Baseline Mean	0.410	0.496	0.859	0.027	0.060	0.116	0.059	0.051	0.008	0.007
R-squared	0.111	0.016	0.118	0.072	0.167	0.109	0.043	0.381	0.123	0.058
School-Grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Specific Linear Time Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. All regressions include grade-year, school-grade fixed effects, as well as a school-specific linear trend. Individual controls include gender, race, FRPL status, and accommodation status. All regressions are weighted by the inverse number of times a student is observed in the sample. Public school voucher students are those that used the voucher to move from the public to private sector. The marginal effect of adding one voucher student to a class of 40 is shown in brackets, and is defined as the coefficient divided by 40. Data on enrollment and test scores come from the IDOE-CREO database. The number of observations in Column (1) are significantly lower because we cannot include prior test score for third graders.

Table 3: Results on Student Achievement

VARIABLES	Std. Test Score	Std. Test Score	Std. Math Score	Std. Math Score	Std. ELA Score	Std. ELA Score
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure to Public School Voucher Students						
Fraction Students using Voucher	-0.243**	-0.200**	-0.261**	-0.239**	-0.178*	-0.131
	(0.107)	(0.094)	(0.122)	(0.108)	(0.106)	(0.103)
	[-0.006]	[-0.005]	[-0.007]	[-0.006]	[-0.004]	[-0.003]
Observations	113,618	113,618	113,618	113,618	113,618	113,618
R-squared	0.189	0.197	0.182	0.191	0.175	0.182
School-Grade FE	Yes	Yes	Yes	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School Specific Linear Time Trend	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. All regressions include grade-year and school-grade fixed effects. Individual controls include gender, race, FRPL status, and accommodation status. All regressions are weighted by the inverse number of times a student is observed in the sample. Public school voucher students are those that used the voucher to move from the public to private sector. The marginal effect of adding one voucher student to a class of 40 is shown in brackets, and is defined as the coefficient divided by 40. Data on enrollment and test scores come from the IDOE-CREO database.

Table 4: Results on Cohort Characteristics

	Percent Female	Percent White	Percent Black	Percent Hispanic	Percent FRPL	Percent ELL	Percent Special Educ.	Total Number Students
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure to Public School Voucher Students								
Fraction Students using Voucher	0.025	-0.256***	0.064***	0.092***	0.279***	0.056*	-0.022	1.582
	(0.040)	(0.037)	(0.015)	(0.027)	(0.038)	(0.030)	(0.018)	(2.339)
Observations	113,618	113,618	113,618	113,618	113,618	113,618	113,618	113,618
Baseline Mean	0.487	0.848	0.023	0.051	0.103	0.025	0.052	39.570
R-squared	0.283	0.848	0.788	0.894	0.798	0.546	0.526	0.933
School-Grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Specific Linear Time Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. All regressions include grade-year, school-grade fixed effects, as well as a school-specific linear trend. Individual controls include gender, race, FRPL status, and accommodation status. All regressions are weighted by the inverse number of times a student is observed in the sample. Public school voucher students are those that used the voucher to move from the public to private sector. The marginal effect of adding one voucher student to a class of 40 is shown in brackets, and is defined as the coefficient divided by 40. Data on enrollment and test scores come from the IDOE-CREO database.

Table 5: Heterogeneity Effects of Voucher Students

VARIABLES	(1) Overall Test Score	(2) Math Test Score	(3) ELA Test Score
Panel A: Identifies as Female			
Fraction Students Using Public School Voucher	-0.158	-0.209*	-0.110
	(0.109)	(0.121)	(0.121)
Interaction with Identifies as Female	-0.096	-0.073	-0.051
	(0.123)	(0.122)	(0.134)
Observations	113,618	113,618	113,618
Panel B: Identifies as non-White			
Fraction Students using Public School Voucher	-0.044	-0.081	0.004
	(0.100)	(0.114)	(0.107)
Interaction with Identifies as non-White	-0.779***	-0.788***	-0.671***
	(0.186)	(0.196)	(0.178)
Observations	113,618	113,618	113,618
Panel C: Prior Test Performance			
Fraction Students using Public School Voucher	0.059	0.116	0.001
	(0.116)	(0.125)	(0.116)
Interaction with Prior Year's Score in Bottom Tercile	-0.406***	-0.573***	-0.278**
	(0.121)	(0.130)	(0.124)
Interaction with Prior Year's Score in Top Tercile	-0.303*	-0.554***	-0.003
	(0.168)	(0.177)	(0.172)
Observations	92,549	92,549	92,549

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. All regressions include grade-year, school-grade fixed effects, as well as a school-specific linear trend. Individual controls include gender, race, FRPL status, and accommodation status. All regressions are weighted by the inverse number of times a student is observed in the sample. Public school voucher students are those that used the voucher to move from the public to private sector. Data on enrollment and test scores come from the IDOE-CREO database.

Table 6: Heterogeneity By Voucher Student Characteristics

VARIABLES	(1) Preferred Model	(2) Bottom Tercile Voucher Students	(3) Top Tercile Voucher Students
Panel A: Overall Test Scores			
Fraction Students Using Public School Voucher	-0.200**	-0.390**	0.154
	(0.094)	(0.166)	(0.246)
	[-0.005]	[-0.100]	[0.004]
R-squared	0.197	0.199	0.199
Observations	113,618	95,407	95,407
Panel B: Math Test Scores			
Fraction Students using Public School Voucher	-0.239**	-0.559***	0.037
	(0.108)	(0.198)	(0.268)
	[-0.006]	[-0.014]	[0.001]
R-squared	0.191	0.193	0.193
Observations	113,618	95,407	95,407
Panel C: ELA Test Scores			
Fraction Students using Public School Voucher	-0.131	-0.369**	0.242
	(0.103)	(0.180)	(0.211)
	[-0.003]	[-0.009]	[0.006]
R-squared	0.182	0.186	0.186
Observations	113,618	95,407	95,407

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. All regressions include grade-year, school-grade fixed effects, as well as a school-specific linear trend. Individual controls include gender, race, FRPL status, and accommodation status. All regressions are weighted by the inverse number of times a student is observed in the sample. Public school voucher students are those that used the voucher to move from the public to private sector. Data on enrollment and test scores come from the IDOE-CREO database.

Table 7: Heterogeneity by Program Expansion

VARIABLES	Test Score Pre-Expansion	Test Score Post-Expansion	Math Score Pre-Expansion	Math Score Post-Expansion	ELA Score Pre-Expansion	ELA Score Post-Expansion
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure to Public School Voucher Students						
Fraction Students Using Voucher	-0.419***	-0.145	-0.490***	-0.103	-0.296*	-0.161
	(0.142)	(0.148)	(0.157)	(0.154)	(0.170)	(0.150)
	[-0.010]	[-0.004]	[-0.012]	[-0.003]	[-0.007]	[-0.004]
Observations	38,430	75,181	38,430	75,181	38,430	75,181
R-squared	0.217	0.199	0.212	0.196	0.205	0.182
School-Grade FE	Yes	Yes	Yes	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School Specific Linear Time Trend	No	No	No	No	No	No

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. All regressions include grade-year and school-grade fixed effects. Individual controls include gender, race, FRPL status, and accommodation status. All regressions are weighted by the inverse number of times a student is observed in the sample. Public school voucher students are those that used the voucher to move from the public to private sector. The marginal effect of adding one voucher student to a class of 40 is shown in brackets, and is defined as the coefficient divided by 40. Data on enrollment and test scores come from the IDOE-CREO database.

Table 8: Descriptive Statistics

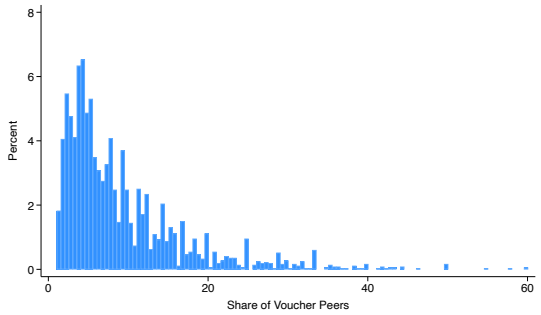
	Voucher Students 2012-2013		Voucher Students 2014 and After		
	Mean	SD	Mean	SD	**Difference**
	(1)	(2)	(3)	(4)	(5)
Panel A: Demographic Characteristics					
White	0.481	(0.500)	0.458	(0.498)	0.023*
Black	0.217	(0.412)	0.194	(0.396)	0.023**
Hispanic	0.205	(0.403)	0.261	(0.439)	-0.056***
Female	0.510	(0.500)	0.517	(0.500)	-0.007
FRPL	0.767	(0.423)	0.753	(0.431)	0.014
Has Math Accommodation	0.139	(0.346)	0.164	(0.371)	-0.025***
Has ELA Accommodation	0.139	(0.346)	0.167	(0.373)	-0.028***
Panel B: Education Outcomes					
Standardized Test Score	-0.238	(0.902)	-0.222	(0.913)	-0.016
Standardized Math Score	-0.244	(0.914)	-0.230	(0.919)	-0.013
Standardized ELA Score	-0.194	(0.911)	-0.179	(0.928)	-0.015

Notes: This table shows descriptive statistics on voucher students that entered participating private school before and after the expansion of the Indiana Choice Scholarship Program. Data on enrollment and test scores come from the IDOE-CREO database. Public school voucher students are those that used the voucher to move from the public to private sector. Column (5) reports the difference in the means between voucher students that entered after 2014 versus in the first two years of the program.

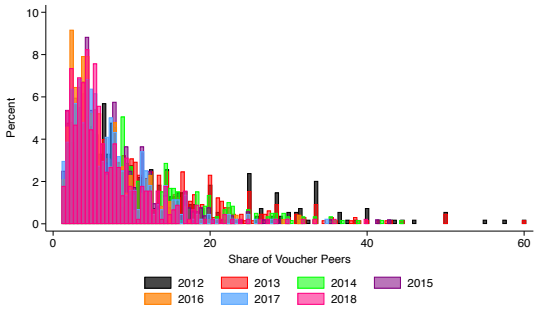
IX Figures

Figure 1: Variation in Share of Voucher Students

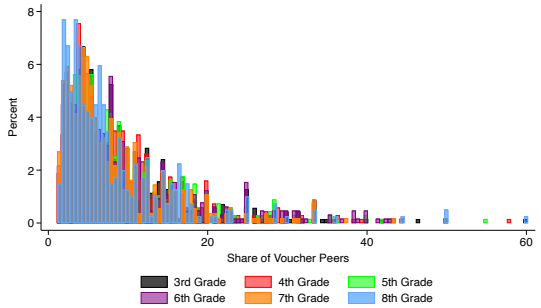
(a) Overall Variation by School-Grade Cohort (>0)



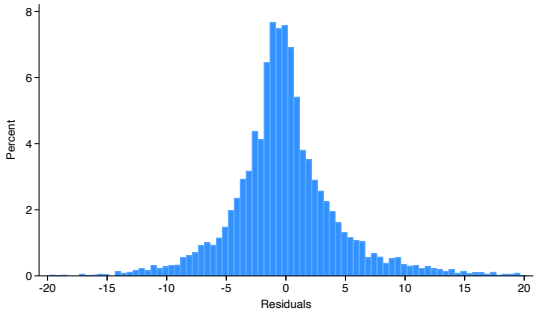
(b) Variation Across Years



(c) Variation Across Grades

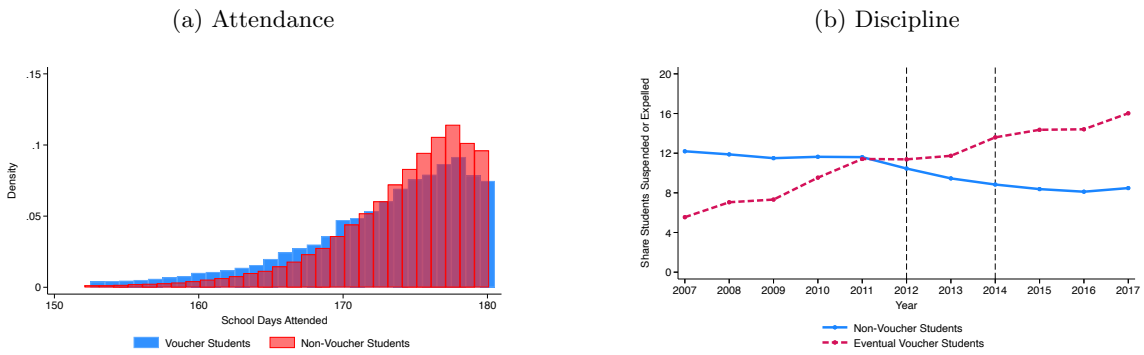


(d) Residualized Variation



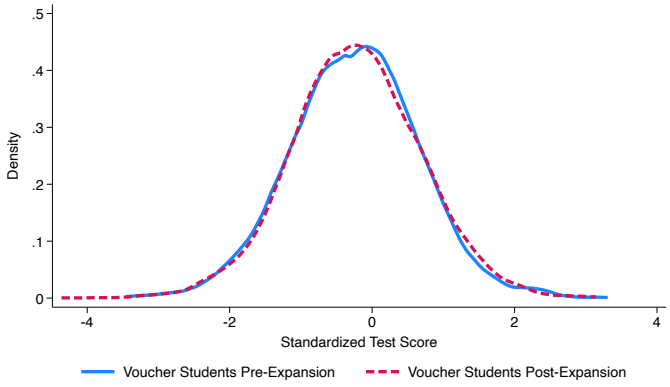
Notes: This figure presents histograms of the variation in voucher students across school-grade cohorts. Panel A presents the variation across all schools and years in the sample period, Panel B plots separate histograms by year, Panel C plots separate histograms by grade, and Panel D residual variation once accounting for the school-grade, grade-year and school-specific linear time trends. Data on enrollment and test scores come from the IDOE-CREO database.

Figure 2: Attendance and Discipline by Voucher Status



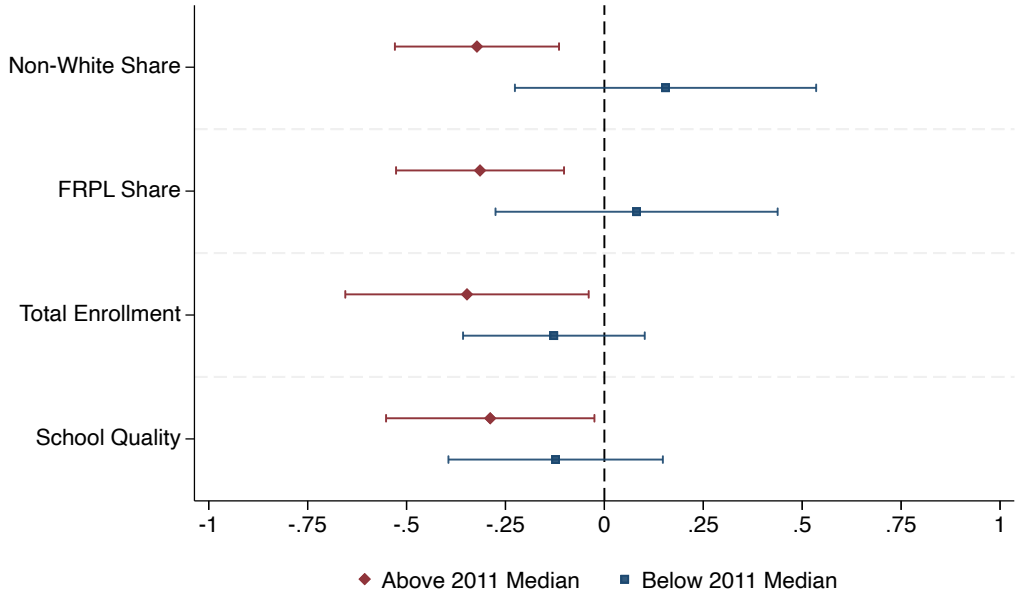
Notes: This figure presents the attendance and discipline rates by voucher status. Panel A presents histograms of the number of school days attended for public school voucher students (blue) and their non-voucher counterparts (red). Public school voucher students are those that used the voucher to move from the public to private sector. Panel B presents average discipline rates by year across public school students that either eventually use a voucher to attend a private school (red) or remain in the public school system (blue). Data on enrollment, attendance and discipline come from the IDOE-CREO database.

Figure 3: Test Score Distribution of Voucher Students by Program Timing



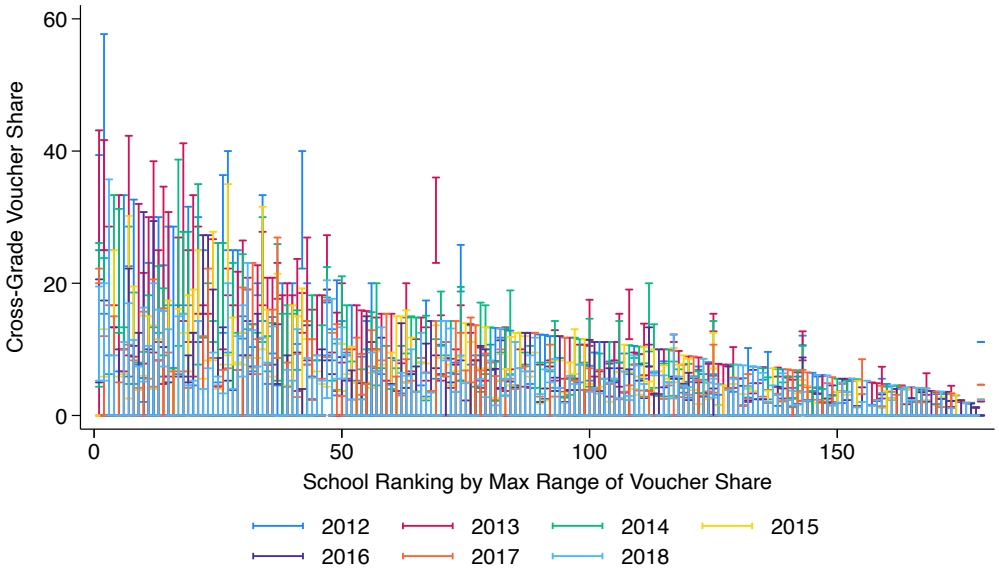
Notes: This figure presents the kernel density plots of standardized test scores for voucher students that attended a participating private school before and after the expansion of the Indiana Choice Scholarship Program. Data on enrollment and test scores come from the IDOE-CREO database.

Figure 4: Heterogeneity by Pre-Policy School Characteristics



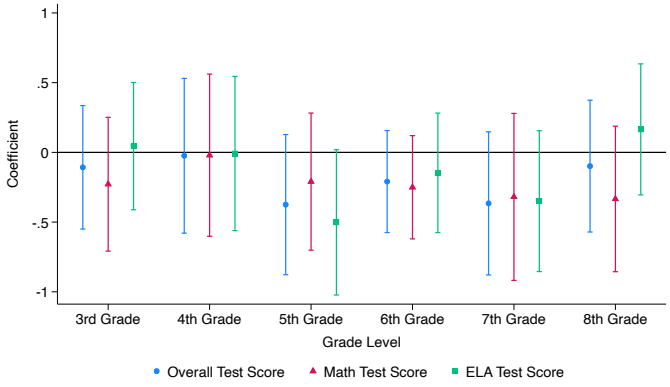
Notes: This figure presents the estimates of Equation 1 when we split the sample by pre-policy school characteristics. Each point is the result of a separate estimation. Red diamonds always represent the results for above median baseline value and Blue squares always represent the results for below median baseline value. 95% confidence intervals are reported. All estimations include school-grade, grade-year fixed effects, as well as school specific linear time trends. Data on enrollment and test scores come from the IDOE-CREO database.

Figure A1: Cross-Grade Variation in Voucher Student Share



Notes: This figure presents the cross-grade variation in the share of voucher students for each school and year in the sample. The bars represent the range of the share of voucher students in school-grade cohort in a given school-year combination. The x-axis orders the schools by the max range of voucher share across the entire sample period. Data on enrollment come from the IDOE-CREO database.

Figure A2: Results by Grade Level



Notes: This figure presents the estimates of Equation 1 when we split the sample by incumbent private school students grade level. Each point is the result of a separate estimation. Blue circles always represent the effects on overall test scores, red triangles always represent the effects on math test scores, and green squares always represents the effects on English language arts (ELA) test scores. All estimations include school-grade, grade-year fixed effects, as well as school specific linear time trends. Data on enrollment and test scores come from the IDOE-CREO database.

Table A1: Stability of Results on Student Achievement

	Overall				Math				ELA			
	No Controls No Weights	Weights Only	Controls Only	Preferred Model	No Controls No Weights	Weights Only	Controls Only	Preferred Model	No Controls No Weights	Weights Only	Controls Only	Preferred Model
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Exposure to Public School Voucher Students												
Fraction Students using Voucher	-0.132	-0.246**	-0.130	-0.200**	-0.152	-0.289**	-0.140	-0.239**	-0.092	-0.165	-0.101	-0.131
	(0.091)	(0.101)	(0.087)	(0.094)	(0.101)	(0.114)	(0.098)	(0.108)	(0.097)	(0.109)	(0.094)	(0.103)
Observations	113,618	113,618	113,618	113,618	113,618	113,618	113,618	113,618	113,618	113,618	113,618	113,618
R-squared	0.104	0.110	0.184	0.197	0.121	0.125	0.182	0.191	0.073	0.082	0.167	0.182
School-Grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Specific Linear Time Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Weights	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. Individual controls include gender, race, FRPL status, and accommodation status. All regressions are weighted by the inverse number of times a student is observed in the sample. Public school voucher students are those that used the voucher to move from the public to private sector. The marginal effect of adding one voucher student to a class of 40 is shown in brackets, and is defined as the coefficient divided by 40. Data on enrollment and test scores come from the IDOE-CREO database.

Table A2: Results on Student Achievement with Expanded Sample

VARIABLES	Std. Test Score	Std. Test Score	Std. Math Score	Std. Math Score	Std. ELA Score	Std. ELA Score
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure to Public School Voucher Students						
Fraction Students using Voucher	-0.195**	-0.172**	-0.190*	-0.182**	-0.157**	-0.139**
	(0.083)	(0.067)	(0.099)	(0.084)	(0.073)	(0.066)
	[-0.005]	[-0.004]	[-0.005]	[-0.005]	[-0.004]	[-0.003]
Observations	169,545	169,545	169,545	169,545	169,545	169,545
R-squared	0.231	0.238	0.216	0.226	0.211	0.217
School-Grade FE	Yes	Yes	Yes	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School Specific Linear Time Trend	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. Individual controls include gender, race, FRPL status, and accommodation status. All regressions are weighted by the inverse number of times a student is observed in the sample. Public school voucher students are those that used the voucher to move from the public to private sector. The marginal effect of adding one voucher student to a class of 40 is shown in brackets, and is defined as the coefficient divided by 40. The sample includes all students in participating private schools that did not use a voucher to move from the public to private sector. Data on enrollment and test scores come from the IDOE-CREO database.

Table A3: Falsification Test Using Non-Voucher New Students

VARIABLES	Std. Test Score	Std. Math Score	Std. ELA Score
	(1)	(2)	(3)
Fraction Non-Voucher New Students	-0.097	-0.125	-0.061
	(0.071)	(0.081)	(0.067)
	[-0.002]	[-0.003]	[-0.002]
Observations	106,827	106,827	106,827
R-squared	0.199	0.194	0.182
School-Grade FE	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes
School Specific Linear Time Trend	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. All regressions include grade-year, school-grade fixed effects, as well as a school-specific linear trend. Individual controls include gender, race, FRPL status, and accommodation status. All regressions are weighted by the inverse number of times a student is observed in the sample. Public school voucher students are those that used the voucher to move from the public to private sector. The marginal effect of adding one non-voucher, new student to a class of 40 is shown in brackets, and is defined as the coefficient divided by 40. Data on enrollment and test scores come from the IDOE-CREO database.

Table A4: Validity Check - Changing Schools and Attrition

VARIABLES	(1) Change Schools by FRPL Status	(2) Change Schools By Race	(3) Missing Test by FRPL Status	(4) Missing Test by Race
Fraction Students using Voucher	-0.108	-0.149	-0.017	0.005
	(0.073)	(0.140)	(0.014)	(0.024)
Interaction with	0.107	-	0.030	-
Qualifies for FRPL	(0.173)	-	(0.034)	-
Interaction with	-	0.059	-	-0.023
Identifies as White	-	(0.143)	-	(0.029)
R-squared	0.123	0.123	0.058	0.058
Observations	74,574	74,574	114,332	114,332
School-Grade FE	Yes	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes	Yes
School Specific Linear Time Trend	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. All regressions include grade-year, school-grade fixed effects, as well as a school-specific linear trend. Individual controls include gender, race, FRPL status, and accommodation status. All regressions are weighted by the inverse number of times a student is observed in the sample. Public school voucher students are those that used the voucher to move from the public to private sector. Data on enrollment and test scores come from the IDOE-CREO database.

Table A5: Comparing Switchers versus Stayers

VARIABLES	Female	White	Black	Hispanic	FRPL	Special Ed.	Section 504	Math Acc.	ELA Acc.	Test Score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Switches School	-0.009	-0.042***	0.013***	0.018***	0.043***	0.017***	0.005**	0.023***	0.023***	-0.204***
	(0.007)	(0.005)	(0.003)	(0.004)	(0.004)	(0.004)	(0.002)	(0.004)	(0.004)	(0.012)
Observations	74,490	74,490	74,490	74,490	74,490	74,490	74,490	74,490	74,490	74,490
Baseline Mean	0.500	0.883	0.0177	0.0464	0.0651	0.0448	0.0156	0.0444	0.0437	0.428
R-squared	0.072	0.174	0.134	0.221	0.197	0.096	0.114	0.104	0.104	0.187
School-Grade-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. All regressions include school-grade-year fixed effects. All regressions are weighted by the inverse number of times a student is observed in the sample. Switches schools is an indicator variable identifying students that no longer attended the private school in the following academic year. The sample only includes students in non-traditional moving grades.

Table A6: Heterogeneity Effects of Voucher Students

	(1)	(2)	(3)
VARIABLES	Overall Test Score	Math Test Score	ELA Test Score
Fraction Students using Voucher	-0.159	-0.045	-0.157
	(0.107)	(0.123)	(0.125)
Interaction with Prior Year's Score in Bottom Tercile	-0.110	-0.253**	-0.001
	(0.109)	(0.124)	(0.130)
Interaction with Prior Year's Score in Top Tercile	0.037	-0.288**	0.240
	(0.146)	(0.145)	(0.171)
Observations	92,549	92,549	92,549
R-squared	0.622	0.591	0.502
School-Grade FE	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes
School Specific Linear Time Trend	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All standard errors are clustered at the school level. Each coefficient is the result of a separate estimation. All regressions include grade-year, school-grade fixed effects, as well as a school-specific linear trend. Individual controls include gender, race, FRPL status, and accommodation status. All regressions are weighted by the inverse number of times a student is observed in the sample. Public school voucher students are those that used the voucher to move from the public to private sector. Data on enrollment and test scores come from the IDOE-CREO database.

B1 Appendix - Qualitative Analysis and Coding Procedures

Interview Sample and Analytic Purpose

The qualitative component of this study used two waves of semi-structured interviews conducted with teachers and incumbent non-voucher students in private schools participating in the Indiana Choice Scholarship Program (ICSP). The first wave was conducted during the 2012-13 school year, in the second school year after the initial implementation of the program, while the second wave was conducted during the 2017-18 and beginning of the 2018-19 school years, after voucher participation had grown substantially. As described in Section III.B of the manuscript, the original interview protocols included questions about instructional and curricular changes, classroom dynamics, student integration, and the experiences of students already enrolled in participating private schools.

We restricted our analysis to interviews with teachers serving in the same grade levels as tested in the quantitative analysis (grades 3-8) and the student sample to incumbent non-voucher students in those same grade levels. The final teacher sample contained 58 interviews, including 33 interviews in the first wave and 25 in the second, including ten of the same teachers. The student sample contained 22 interviews, including ten in the first wave and 12 in the second. See Table B1 for a full count of the interviews coded.

The interviews are not used for identification or causal estimation as part of this paper. Instead, we use them to examine whether teachers and students described classroom processes that are consistent with potential mechanisms underlying the quantitative findings. The interviews were collected well before development of the quantitative analysis in this paper; therefore, they were not explicitly designed in response to the quantitative results. However, the interviews were contemporaneous with our quantitative analytical window, enabling us to use a sequential explanatory analysis approach with secondary data (Creswell and Plano Clark, 2018).

Development of the Deductive Explanatory Coding Framework

We developed separate deductive coding frameworks and domains for the teacher and student interviews. Both the quantitative analysis in this paper and prior research using these interviews helped to inform these frameworks. We developed and refined the coding frameworks with a pilot set of five teacher and five student interviews before moving forward to the final analysis. Specifically, the pilot coding was used to clarify boundaries between domains, tighten inclusion and exclusion criteria, and reduce over-inclusion of coded turns of text.

The teacher framework initially focused on behavior and classroom management, differentiation, instructional time, and broader classroom disruption and was refined through pilot coding. After

these refinements, the teacher framework distinguished underlying classroom conditions from instructional responses and produced five teacher domains: instructional attention, classroom composition and heterogeneity, behavioral or classroom management, instructional practice and curriculum, and teacher workload or strain.

The student framework was initially developed and further refined around five corresponding domains of the student experience: academic experience, instructional attention, classroom composition and climate, social integration, and voucher awareness. The coding framework for instructional attention further distinguished between general access to teacher assistance and explicit competition for teacher attention, while the social integration framework distinguished peer integration difficulties from broader peer-relationship processes. We excluded statements about academic performance alone (e.g., students reporting their grades) unless they substantively described learning conditions that influenced their academic performance. With the final teacher and student coding frameworks, we moved to coding the full set of interviews.

Dual-Layer Coding Procedure

Each interviewee's complete turn of speech served as the unit of coding, and we retained interviewee responses as complete turns rather than artificially dividing them into smaller excerpts. This preserves the context in which respondents described classroom processes and reduces any tendency to over apply or inflate the number of codes.

We applied two complementary layers of deductive coding to each interviewee turn. The first, mechanism-focused layer, is deliberately restrictive. A turn qualified for code application when it contained at least one of three features: (1) a temporal change or comparison; (2) a comparison across students or groups, such as new versus incumbent students; or (3) a process connecting classroom conditions to instruction, classroom experience, or student learning. We applied codes in instances where interviewees provided limited contextual inference when such a process was strongly implied even without an explicit before-and-after statement.

The second, general experience layer, identified substantive references to the same underlying domains without requiring evidence that the respondent perceived the condition as new or changing. This distinction is particularly important for comparisons across interview waves. Respondents interviewed shortly after implementation were more likely to describe new enrollment and classroom conditions relative to their prior experiences. By the second wave, teachers and students had experienced expanded voucher participation for several years and could describe similar conditions as routine rather than as changes. The two layers therefore allow us to distinguish between evidence about implementation-related mechanisms and the continuing prevalence of classroom conditions

that may have become normalized over time.

In some instances, we assigned multiple codes to the same turn when the content within the turn independently met the criteria for inclusion for each domain. We summarized the resulting classifications for teachers and students at the interview level, reporting the share of interviews containing at least one qualifying turn within each domain. Because teacher interviews were longer in general and the turns themselves contained more rich information, we also reported the share of total teacher turns containing qualifying turns within each domain. These frequencies are descriptive and should not interpreted as estimates representative of all Indiana Choice Scholarship Program-participating private schools at the time of the interviews.

AI-Assisted Coding and Researcher Validation

We used Generative AI (ChatGPT Model 5.6 from OpenAI, within an institutional ChatGPT EDU workspace) as a complementary and interactive research tool throughout the qualitative analysis process. All interviews analyzed by ChatGPT were de-identified before any analysis and our research protocol was approved through the University of Notre Dame's IRB and Human Subjects Research through protocol 26-08-10301. First, we used AI in refining the coding framework through looking at a set of pilot interviews as a companion to initial researcher-led coding. We also compared the researcher-led pilot coding with earlier human coding of the same interviews conducted as part of the original ICSP research project. This comparison served as a validation exercise rather than an attempt to reproduce the earlier thematic framework, which was developed for different research questions. In the development of the final coding framework, the substantive domains, analytic objectives, unit-of-analysis rules, and initial inclusion and exclusion criteria were specified by the research team rather than independently generated by the AI model. The final coding framework intentionally prioritized specificity over exhaustive coverage.

We then used the same AI tool to systematically apply the deductive coding rules to turns across the interviews. The full coding process was iterative, resulting in multiple instances of repeating coding procedures after researcher review identified areas requiring refinement. One author led the qualitative analysis, reviewed intermediate coding outputs, identified classifications that did not adequately represent the intended constructs, refined coding instructions and decision rules, and reviewed the resulting evidence. In compiling the final evidence and results contained herein, the research team retained responsibility for the validation of classifications supporting substantive claims, selection and verification of illustrative quotations, calculation of code prevalence, and interpretation of the qualitative evidence.

Table B1: Qualitative Analytic Sample

	Wave 1	Wave 2
Teacher interviews coded	33	25
Student interviews coded	10	12
Total coded interviews	43	37
Teacher interviews repeated across waves	-	10

Notes: This table reports the counts of the number of interviews coded across the two waves. The sample includes semi-structured interviews with teachers and incumbent non-voucher students in grades 3-8 at eight participating private schools. Wave 1 interviews were conducted in the 2012-13 academic year. Wave 2 interviews were conducted in the 2017-18 and early 2018-19 academic years.

Table B2: Teacher Deductive Coding Framework

Domain	**Domain Description**	**Mechanism Focused Inclusion**	**General Experience Inclusion**	**Exclusions**
Instructional Attention	Allocation of teacher time and individualized instructional support	Changes in differentiation, reteaching, grouping, or allocation of attention associated with student needs/composition	Substantive discussion of differentiation, grouping, reteaching, or individualized attention	Generic statements about helping students
Classroom Composition / Heterogeneity	Variation in student preparedness, ability, background, or prior schooling	Change/comparison involving student mix, preparedness, ability range, or new entrants	Substantive description of variation in student composition or readiness	Generic references to students without compositional content
Behavior / Classroom Management	Behavioral differences, discipline, disruption, routines, or classroom control	New/changing management demands or processes affecting classroom conditions or instruction	Substantive discussion of behavior or classroom management	Isolated behavior unrelated to classroom processes
Instructional Practice / Curriculum	Pacing, curricular coverage, grouping, remediation, or instructional strategy	Change in instructional practice or pacing associated with student needs/composition	Substantive discussion of instructional practice, pacing, or curriculum	Routine curricular description without relevant instructional content
Teacher Workload / Strain	Planning, instructional, management, or emotional demands	Increased workload/strain connected to changing students or classroom conditions	Substantive discussion of workload or strain	Generic job demands unrelated to classroom experience

Notes: This table reports the deductive coding framework applied to teacher interview transcripts. Coding frameworks were applied to interviews with teachers in grades 3-8 at participating private schools. The unit of coding is the interviewee's complete turn of speech. Mechanism-focused inclusion requires a temporal change or comparison, a comparison across students or groups, or a process connecting classroom conditions to instruction, classroom experience, or student learning. General experience inclusion captures substantive references to the domain without evidence that the respondent perceived the condition as new or changing. A single turn may receive multiple domain codes.

Table B3: Student Deductive Coding Framework

Domain	**Domain Description**	**Mechanism Focused Inclusion**	**General Experience Inclusion**	**Exclusions**
Academic Experience	Learning conditions, rigor, difficulty, or instructional pace	Change/comparison/ process affecting learning conditions or pace	Substantive discussion of learning conditions or academic experience	Grades, generic difficulty, or self-description as a good/bad student
Instructional Attention: General Teacher Access	Availability and distribution of teacher assistance	Change/comparison in access to teacher support	Substantive description of access to teacher help	Generic praise of teachers
Instructional Attention: Competition for Teacher Access	Displacement or competition for teacher attention	Waiting longer, reduced attention, or teacher occupied with other students	N/A	General teacher support without displacement
Classroom Composition / Climate	New students, class composition, behavior, routines, or classroom environment	Change/comparison/ process associated with composition or new students	Substantive description of classroom composition or climate	General reference to classmates
Social Integration: Difficulty	Friction entering or maintaining peer relationships	Exclusion, difficulty fitting in, friendship disruption, or integration friction	Substantive discussion of integration difficulty	Generic friendship statements
Social Integration: Peer Relationships	Friendship, grouping, inclusion, and peer processes	Changes/comparisons in peer relationships	Substantive discussion of peer relationships	General naming of friends
Voucher Awareness	Awareness of vouchers/ICSP/ public tuition assistance	Explicit recognition of voucher status/program	N/A	References to new students without program awareness

Notes: This table reports the deductive coding framework applied to student interview transcripts. Coding frameworks were applied to interviews with incumbent non-voucher students in grades 3-8 at participating private schools. The unit of coding is the interviewee's complete turn of speech. Mechanism-focused and general experience inclusion criteria are defined as in Table B2. N/A indicates that no general experience layer was applied to the domain. A single turn may receive multiple domain codes.

Table B4: Teacher Interview Code Prevalence

Domain	**Percentage of Wave 1 Interviews with Mechanism**	**Percentage of Wave 2 Interviews with Mechanism**	**Percentage of Wave 1 Turns with Mechanism**	**Percentage of Wave 2 Turns with Mechanism**
Instructional Attention	90.9%	88.0%	4.5%	2.9%
Classroom Composition / Heterogeneity	100.0%	100.0%	9.1%	6.1%
Behavior / Classroom Management	90.9%	92.0%	6.8%	2.8%
Instructional Practice / Curriculum	93.9%	96.0%	7.2%	5.8%
Teacher Workload / Strain	81.8%	72.0%	3.3%	1.6%

Notes: Wave 1 interviews ($N = 33$) were conducted in the 2012-13 academic year; Wave 2 interviews ($N = 25$) were conducted in the 2017-18 and early 2018-19 academic years. Ten teachers were interviewed in both waves. Interview-level columns report the share of interviews containing at least one turn meeting the mechanism-focused criteria for the domain. Turn-level columns report the share of all teacher turns in the wave that meet the mechanism-focused criteria for the domain. Because turns may receive multiple codes, percentages do not sum to 100. Domains are defined in Table B2. Frequencies are descriptive and not representative of all participating private schools.

Table B5: Student Interview Code Prevalence

Domain	**Percentage of Wave 1 Interviews with Mechanism**	**Percentage of Wave 2 Interviews with Mechanism**
Academic Experience	20.0%	0.0%
Instructional Attention: General Teacher Access	30.0%	25.0%
Instructional Attention: Competition for Teacher Access	10.0%	0.0%
Classroom Composition / Climate	90.0%	33.0%
Social Integration: Difficulty	60.0%	33.0%
Social Integration: Peer Relationships	40.0%	17.0%
Voucher Awareness	60.0%	42.0%

Notes: Wave 1 interviews ($N = 10$) were conducted in the 2012-13 academic year; Wave 2 interviews ($N = 12$) were conducted in the 2017-18 and early 2018–19 academic years. Columns report the share of interviews containing at least one turn meeting the mechanism-focused criteria for the domain. Because turns may receive multiple codes, percentages do not sum to 100. Domains are defined in Table B3. Frequencies are descriptive and not representative of all participating private schools.