Supplemental Appendices to "School Proximity, Attendance, Stability, and Achievement among Homeless Students"

Mike Cassidy
Princeton University

October 30, 2023

A Data and Sample

A major contribution of this paper is the construction of an original dataset comprehensively describing student homelessness in New York City. Given NYC's outsized importance in the realm of family homelessness—along with the extensive detail of linked longitudinal administrative data—this represents perhaps the richest portrait of student homelessness in the U.S. to date. The data comes from two key sources: the NYC Department of Homeless Services (DHS) and the NYC Department of Education (DOE). In this section, I summarize key data management steps. Additional information about the DHS data can also be found in Cassidy (2020).

A.1 DHS Data

The DHS portion of the data constitutes the core sample: all eligible families with children entering shelter from January 1, 2010 to December 31, 2016.[1] These records, which contain details about families' compositions, demographics, and conditions of shelter entry, as well as basic identifying information, are extracted from DHS' Client Assistance and Rehousing Enterprise System (CARES), which is the City's management information system for homeless services. This "sample" is essentially a census, excluding only those (rare) individuals with missing data on critical identifying variables.

CARES contains individual-level records for each family member. In Cassidy (2020), I rework this data so that the unit of observation becomes the family-spell. That is, there is one observation per family per shelter stay, with new spells defined as those occurring more than 30 days subsequent to the end of a previous stay.[2]

The DHS data contains rich information about families and their shelter stays, most of which comes from the Temporary Housing Assistance (THA) applications families fill out to apply for shelter. Variables include basic identifying information (e.g., name, date of birth), family relationships, the presence of health issues (physical, mental, or substance abuse), official shelter eligibility reason, and housing history (most recent address prior to application). Shelter stay attributes, including facility type, address, and dates of stay, come from "Lodge History" extracts, a separate CARES query.

The data is collected primarily for management rather than analysis, and so requires extensive processing to be econometrically coherent. Key data management steps include: defining shelter spells (including length of stay calculations), geocoding addresses, and defin-

[1] Due do a minor coding issue that does not affect results, 16 students in this sample potentially had their applications entered or approved outside the calendar year 2010–2016 period.

[2] DHS considers returns to shelter within 30 days of leaving to be part of the same spell.

ing analytical variables (e.g., discrete variables summarizing eligibility reasons or counting family members), some of which are assembled across observations.

I augment this core DHS data by linking it to administrative records maintained by other agencies. I obtain information on public benefit use—Cash Assistance (i.e., CA, public assistance or "welfare," consisting of federal Temporary Assistance for Needy Families (TANF) and New York State (NYS) Safety Net Assistance (SNA)) and the Supplemental Nutrition Assistance Program (i.e., SNAP or "Food Stamps")—from the Department of Social Services (DSS), using probabilistic matching techniques based on Social Security Number (SSN), first name, last name, and date of birth. The DSS data also includes information on race and self-reported educational attainment. In a similar fashion, the NYS Department of Labor (DOL) provides data on quarterly employment and earnings, through a deterministic match on SSN.[3] To ease computational burden, the public benefit and labor matches are restricted to head of family. Because (a) most homeless families consist of a single adult and several children and (b) heads of case are most likely to appear in the benefit and labor data, this restriction should not meaningfully change the results relative to an exhaustive match of all family members.

For purposes of assessing family outcomes, as in Cassidy (2020), the natural unit of observation is the family-spell. In the present study, the underlying individual-level records come to the fore. From the CY2010-2016 CARES census, I cull the records of all individuals aged 4 to 21 during any point in their shelter stays. I choose these cutpoints because they represent the minimum (children can begin pre-K at age 4) and maximum (children can attend school through the school year in which they turn 21) ages individuals can be enrolled in DOE.[4] In total, there are 89,337 unique such children. Using CARES' individual and family identifiers, I then relink these individuals to the family-shelter-spells of which they are a part. In this manner, the unit of observation becomes the individual-homeless-spell.

Several comments are in order regarding the definition of DHS-derived analytical variables. All covariates (see Tables A.4A–A.4B) are defined at the time of shelter entry (or as near as is possible). Person-specific variables, such as age, are defined at the individual level. Correspondingly, attributes shared by all family members, such as eligibility reason or shelter type, are defined at the family level. The exceptions are variables derived from DSS and DOL: CA, SNAP, employment, earnings, and level of education, which are defined by head of household but treated as "family-level" variables common to all members. Families that are not matched to DSS or DOL are assumed genuinely not receiving benefits or not

[3]For simplicity, I refer to the DSS/HRA data and DOL under the umbrella of "DHS" since the linkage is performed with the DHS data.

[4]21-years-of-age is also the DHS definition of child.

employed, respectively (though, due to the fuzzy nature of the match, there may be some false negatives). I take the extra step of creating an "unknown" education category for families that do not match DSS in order to include education as a covariate without restricting the sample. For a similar reason—avoiding unduly excluding families from the sample—I also create an "unknown" category for homeless eligibility reason, which is a DHS CARES variable missing for a handful of families. For earnings, I add one dollar before taking the natural log.

In sum, the DHS data consists of unique observations for each school-age child during each homeless spell experienced by their families, complete with all covariates, both individual and family-level, associated with each spell.

A.2 DOE Data

The second key data source are educational records maintained by DOE, spanning school years 2005-06 to 2016-17. It contains records for each student during each school year, with separate topical tables for June biographical information (demographics, student characteristics, and enrollment details, including school ID and attendance), test scores (3–8 grade state standardized tests and Regents for high schoolers), and graduation (for high schoolers). The biographical table is given the "June" designation because it is reconciled at the end of each school year, in June, and reflects each student's most up-to-date information as of then. For data size reasons—each table includes all public school students, not only homeless ones—there is a separate topical table for each school year. All variables are student-specific.

In addition to the topical tables, there is also a separate Transactions table detailing all admissions and discharges, including normative promotions (i.e., structural leveling up, as from elementary to middle school) as well as nonstructural school changes, during all school years in the sample. In contrast to the June biographical information, the Transactions table records the precise dates and reasons for each school change.

In practice, a fair amount of data processing must take place to shape the records into a form suitable for analysis. Key tasks include harmonizing variables across years (as available fields and definitions change over time) and linking a student's records (via unique student ID's) across topics and over time. The resulting dataset consists of a single observation for each student in each school year.

Key DOE variables used in the analysis are described in Section 3 and Tables A.4A–A.4B. As with the DHS data, some of these variables are not native to the administrative data, but rather are constructed from the underlying fields (e.g., test scores normalized by grade and year).

Of particular note, schools are identified by unique "DBNs," comprised of school district (D), school borough (B), and school number (N) codes. To facilitate measurement of school-shelter distances, I link these DBNs to publicly available school geocode files (covering school years 2012 to 2017), which, in addition to school names and addresses, contain X-Y coordinates as well as grade ranges (which are important for identifying nonstructural school changes).

A.3 Data Match

The matching procedure to link DHS' candidate homeless students with DOE records, performed with the assistance of the Mayor Office's Center for Innovation through Data Intelligence (CIDI) and DOE staff, follows standard City protocols for linking human service and education data. It uses The Link King version 9.0 (Campbell, 2018), a SAS application, with default settings and match records based on first name, last name, date of birth, and sex. The Link King uses a variety of sophisticated algorithms to deterministically and probabilistically match records across datasets. For details, see Campbell, Deck and Cox (2018). I accept match certainty levels 1 (highest possible) to 6 (low-moderate) as true matches, while level 7 (probabilistic maybe), along with unmatched records, are defined as non-matches. Close cases, including those with several match candidates, are reviewed manually. Once the match is complete, data is deidentified by stripping names and official identifiers and replacing them with scrambled student ID.

Given 12 years of education records and 7 years of homeless data, the match is over-inclusive. There are three types of matched students: (1) children who are in school during their shelter stays, (2) adult family members (typically heads of household) who attended DOE schools at some time in the recent past, (3) children too young to be in school during their time in shelter but who enrolled in DOE subsequently. Because I am interested in the contemporaneous and short-term effects of shelter policy, my interest is in the first group.

Even restricting the match sample to age-relevant individuals, the panel nature of the data guarantees a number of irrelevant matches. A non-trivial share of household heads age 18–21 (group 2 above) are, in fact, heads of household who previously completed their DOE careers (given that DOE records extend back to 2005-06). Thus, I trim the match sample by eliminating all matches involving heads of household. Note that, by design, this also excludes all in-school heads of household, on grounds that my primary interest is in outcomes among minor students; adult students with dependents can reasonably be expected to be subject to different, potentially confounding, dynamics. In a similar way, I drop all matches where a homeless child in question is too young to be in school during a homeless spell (group 3

above); these children match due to enrollment in DOE during a subsequent post-shelter year.

Table A.6 details the match results by birth year, focusing specifically on children aged 5–18 during a shelter stay. Overall, 64,728 of 74,058 unique candidate students (87 percent), accounting for 78,465 of 88,582 student-homeless-episodes present in the DHS data (89 percent), have successful DOE matches. For students in the "core" birth (calendar) years of 1995–2008, the match rate is 90 percent or greater; these children are in their prime schooling years during the 2010–16 period that comprises the homelessness window. As expected, match rates are lower for older and younger children.[5]

The upshot of this considerable data processing effort is an unprecedented chronicle of student homelessness, detailing students' educational histories in the context of their families' homelessness experiences, as well as their characteristics, composition, labor market experiences, and public benefit use. I discuss the key variables used in the analysis in Section 3.

For most of the paper, I focus a subset of these observations in order to sharpen the analysis. These sample refinements are summarized in Table A.5 and discussed in Section 3.

A.4 Complete Sample

Beyond the core dataset of homeless students, I also create a second broader dataset inclusive of all students in all available school years. I refer to this as the "complete" data. As shown in Table A.7, it spans school years 2010–2015 and contains 6,798,801 student-school-year observations across 2,177 unique schools.

The purpose of the complete data is to compare homeless students with their housed peers, which provides a frame for interpreting results. To do so, I impose the sequential sample restrictions given by row in Table A.7, arriving at a "complete sample" defined analogously to my main analytical sample and consisting of 4 million student-school-years. Because the homelessness data spans CY2010–2016 shelter entries, students who entered shelter prior to CY2010, and remained in shelter in subsequent school years, are not identified as homeless. This will cause some degree of attenuation bias in housed-homeless contrasts, particularly in the early years of the data. However, because the average length of a homeless spell for students in my sample is just over a year, comparisons from 2011 on should be mostly unaffected.

I also use the complete data to construct school-level covariates for the main analysis.

[5]There are several legitimate reasons a school-age homeless child may not show up in DOE records, including moves into and out of NYC contemporaneous with homeless episodes and enrollment in parochial or private schools.

With the exception of normalized test scores (for which I used the 4 million student-school-years in the complete sample as the reference population), I do not impose sample restrictions when defining school-level covariates so as to capture the most comprehensive school-level characteristics.

A.5 Geographical Data Management

Because geography is central to this paper, it is worth highlighting the software that facilitates the analysis. After cleaning and standardizing addresses, I geocode shelter addresses and families' pre-shelter residential ("home") addresses using the New York City Department of City Planning's Geosupport Desktop Edition (Version 17.1) (NYC Department of City Planning, 2017).

Shelter addresses come from a pair of DHS CARES "Facilities" queries that together detail shelter addresses (in the case of multi-site shelters, at the individual shelter unit level), as well as daily-level capacity and occupancy.

School addresses come from the NYC Department of Education's annual school location geocode files for school years 2012-13 to 2017-18, which are publicly available through NYC Open Data (NYC Department of Education, 2017).

Home-to-shelter and home-to-school distances, as well as linear school-shelter distance, are calculated with the Pythagorean theorem, using geocoded X-Y coordinates.[6] School-to-shelter distances—my primary treatment measure—are calculated using Google Maps Platform's Distance Matrix API (Google, 2023).

A.6 Treatment Definition Details

While the DHS data contains exact dates of shelter stay, DOE's preferred source of school enrollment, the June Biographical data, reports only students' end-of-year status. Thus, using this field is inappropriate for assigning schools of origin.[7]

To address this concern, I turn to the DOE Transactions data and employ the following algorithm to identify each student's original school for each school year.

If a student's first school year is present in the data, they are assigned the school of their first-ever DOE admission from the transactions data for this school year. Students who entered DOE prior to 2005 are assigned their June 2005-06 school.

[6]The GBAT geocoding application establishes Cartesian coordinates according to the State Plane Coordinate system, New York-Long Island zone, NAD 83, with units measured in feet.

[7]In addition, about 10 percent of homeless students originate from outside NYC during the 2010–2015 school years. I exclude these non-NYC students from the analysis.

Next, students with "structural" school changes—that is, scheduled promotion into middle school (usually grade 6) or high school (usually grade 9)—are assigned the school of first transaction for that school year.

For all remaining school years (those which are neither a student's first in DOE nor entail structural changes), students are assigned the school borough of the prior June (i.e., the default assumption is school stability). If prior year school is missing, they are assigned the school of first admission in the current school year; if transactions records are also missing, they are assigned the end-of-year June school.

By assigning each student the earliest possible school with which they are associated in each school year, the risk of mechanical treatment is minimized.

A second issue is that, while the school-shelter nexus is the most policy relevant treatment definition—the explicit goal, after all, is to keep children in their "home" schools—it is not the only sensible way to define treatment. For each student, there are three relevant locations: home (pre-shelter residence), school, and shelter. A function of any of these three pairwise links can identify a coherent treatment concept.

I choose to focus of the school-shelter link for two reasons. First, as proxies for genuine "home" locations, school identities are likely to be more stable and less error prone than address of prior residence, as the latter is both self-reported and more transient, given frequent moves among families at-risk of homelessness. Second, the main interest of this paper is the effect of shelter proximity on educational outcomes, so the relevant distance is that between shelter and school, regardless of prior residence. In practice, there is substantial overlap between the treatment concepts.

Finally, I define treatment at the level of the individual student, rather than for the family as a whole. Although the official policy considers an entire family treated if it is placed in the borough of its youngest child's school, older siblings do not necessarily attend schools in the same boroughs. Untreated students in "treated" families will dilute the effects of proximity, so I focus on the personal measure. In practice, it is rare for siblings to have very different commutes.

B Supplementary Results

B.1 Assessing the Parallel Trends Assumption of Difference-in-Differences

The credibility of DID rests fundamentally on the parallel trends assumption, so Figure A.14 assesses pre-trends in student outcomes for the subsample of K–8 graders observed at least

three years prior to shelter entry. I obtain point estimates and pointwise standard errors (indicated by bars) by running event study regressions of the following form:

$$Y_{it} = \alpha_i + \phi_r + \sum_{\substack{r \in \\ \{-2,-1,1\}}} \Big(\mathbf{1}\{R_{ir} = r\} \times \mathbf{1}\{D_{i1}^B = 1\} \Big) \rho_r + \upsilon_{ir} \tag{1}$$

where outcome Y for student i in year t is function of individual fixed effects (α_i), time fixed effects (ϕ_r), and interaction between the indicator for out-of-borough treatment group membership $D_{i1}^B = 1$ and a set of indicators for time relative to the year prior to shelter entry, $R_{it} = t - (E_i + 1)$, where E_i denotes i's school year of shelter entry, and period $r = 0$ (the year immediately preceding shelter) entry is the omitted baseline. Standard errors, υ_{it}, are clustered by family group. Following best practice (Roth et al., 2023; Rambachan and Roth, 2023; Freyaldenhoven et al., 2021), I also estimate uniform confidence bands (denoted by spikes) and and report the Rambachan and Roth (2023) $\overline{M}$ statistic, which gives the multiple by which post-treatment violation of parallel trends would have to exceed the pre-treatment maximum for the results to be invalidated.

In short, there are no (even pointwise) statistically significant pre-trends, but—to preview main DID results—there are clear increases in absences and school changes among students placed in shelters outside of their school boroughs. The post-treatment estimates for these outcomes are large enough to be tolerant of parallel trends violations double that of the largest pre-treatment violation. Test scores and promotion exhibit neither pre- nor post-trends. Figure A.15 provides a complementary analysis for distance treatment.

B.2 The Natural Experiment and Student Fixed Effects Robustness

Table A.9 assesses commute distance treatment coefficient stability across a range of covariate specifications, starting with no covariates and building up to the school and shelter FE specification. Covariates have little impact on the estimates, lending additional support to the quasi-random shelter assignment assumption.

Table A.10 considers several alternative treatment measures: linear distance (Column 2), commute distance with the walk-transit threshold reduced to 0.5 miles (Column 3), transit distance exclusively (Column 4), walk distance exclusively (Column 5), and commute time (in minutes) (Column 6). Aside from innocuous scale effects (e.g., the commute distance mean is 50 percent larger than the linear distance mean), the results are not sensitive to treatment definition. As might be expected, at treatment means, commute time has a larger

impact on student outcomes than does distance. Compared to hypothetical students with no commutes, students with 46-minute commutes are: absent 2.4 more days, 14.6 pp more likely to change schools, and score 0.03σ less in math.

Table A.11 tests for selection effects by checking the sensitivity of results to alternative sample restrictions. In particular, one might worry that since families are neither required to accept shelter placements nor to remain in them, families with unfavorable placements may leave sooner and/or try to negotiate better assignments. To address these concerns, Column 2 limits the sample to students remaining in their initial shelter assignments (to guard against within-shelter moves)[8]; Column 3 limits the sample to students remaining in shelter at least 30 days (to guard against abbreviated spells); Column 4 keeps each student's first-observed spell only (to guard against "shopping around"); Column 5 includes outliers whose commute distances are at or above the 95th percentile; and Column 6 includes all of the preceding sample modifications at once. Comfortingly, results are generally similar or larger (in absolute value) in the alternative samples.

Table A.12 repeats Table 1 for several alternative definitions of proficiency among 3–8 graders. Rows 1–3 use a binary measure that defines proficient as scoring level 3 or 4 and not missing the test. The next three rows measure proficiency only among test takers, omitting students who miss tests. Rows 7 and 8 use numeric "scale" scores. The results are consistent with the main analysis. Overall proficiency decreases by about 10 percent with borough-level distance changes.

B.3 Additional Heterogeneity Analysis

It seems clear that homeless students benefit from proximity, on average. But the methods and models employed in the main analysis are restrictive in the sense that they estimate a single parameter for an entire population. Given scarcity—not every student can be assigned to their first-best school (or to their first-best place of residence)—it is important to understand who (if anyone) benefits most from proximity to school and why. Tables 3 and A.14A–A.14E conduct a thorough heterogeneity analysis by repeating the main OLS estimating equation (Table 1, Column 2) separately for subgroups split by characteristics of interest. In these tables, columns index outcomes, horizontal panels denote characteristics, and individual rows divide the sample into subgroups defined by levels of the panel's characteristic. As before, each cell gives the coefficient on commute distance, along with the clustered standard error and sample size. The section for each characteristic also contains a formal test for differences in mean treatment effects across levels, and reports these dif-

[8]Specifically, to students whose shelter of longest residence in a particular spell matches their initial shelter assignment.

ferences, as well as associated standard errors and p-values. For binary characteristics, the comparison is between the two levels; for ordered categorical variables, the comparison is between the highest and lowest levels ("unknowns" are excluded).

Table 3 is discussed in the main text. Tables A.14A–A.14D explore how treatment effects vary (or not) among a variety of other student and family characteristics. There is essentially no statistically meaningful heterogeneity among across these characteristics.

In contrast, Table A.14E provides some evidence that there is heterogeneity according to which sorts of schools students attend. In particular, attendance treatment effects are larger (0.214 days/mile vs. 0.137 days/mile) for students whose schools of origin have above median absences (as measured in the complete sample of homeless and housed K–8 graders). However, this finding is not entirely robust: the magnitude is qualitatively similar, but imprecisely measured when the sample is split according to absence rates and shares of students changing schools. On the other hand, there is also evidence that students switching into worse schools (as measured by quartile of the change in standardized days absent from beginning-to-end-of-year schools) as more elastic with respect to distance. Students in the fourth (worst) quartile of absence changes see absences increase by 0.238 days/mile, compared to an (imprecise) 0.089 days/mile among first quartile students. In tandem, Table A.15 shows that when homeless students change schools, they, on average, change into schools that are modestly worse—but this effect decreases with distance. In other words, the further is a student's placement from their school of origin, the better the chance they have to switch into a (somewhat) better school.

Fittingly for a placed-based policy, a final dimension of heterogeneity to consider is the geographic distribution of treatment effects. Figures A.16—A.19 depict the results of subsample regressions for days absent, school change, and math and English test scores, respectively, estimated individually within each school district.

The geographic pattern differs by outcome. The largest effects of commute distance on attendance occur in remote boroughs (Staten Island and the Bronx) or in the remote parts of boroughs (South Brooklyn, Northeast Manhattan, and Eastern Queens). For school changes, geographically isolated Staten Island remains a high responder, but the other places where school change effects are most pronounced are core areas of Brooklyn, Manhattan, and Queens. Similarly, math (and to a lesser extent, English) test scores appear most impacted in places where attendance effects are small (lower Manhattan, Queens, and mid-Brooklyn).

One explanation for this discrepancy is that absences and school changes may, in some sense, be substitutes. Students who opt change schools upon entering shelter may miss fewer days and consequently experience better performance. However, it is important not to read too closely into the geographic results, as effect estimates are not especially precise in many

school districts.

B.4 Mechanisms Analysis

Having shed some light on who benefits (everyone, but especially those situated very near their schools), another key question is why—what are some of the channels through which proximity effects operate?

One potentially important channel is through the mechanism of school change. Changing schools is the outcome most affected by proximity, so it is reasonable to ask whether school change mediates other observed outcomes. Of course, changing schools is also potentially endogenous, so any mediating correlations mobility may have are best interpreted as associations.

Table A.16 provides this analysis. It repeats the main OLS specification (Table 1, Column 2), but interacts an indicator for school change with commute distance. Columns index outcomes, and the rows list the main and interaction effects of commute distance and school change.[9] While these results should not be given causal interpretation (a decision to change schools is not random), they are suggestive. As before, longer commutes are associated with worse outcomes (an additional 0.21 absences/mile). Unsurprisingly, school change (and the disruption that comes along with it) are significantly associated with adverse outcomes—an increase in 2.7 days absent and drops of 0.08σ in math and 0.03σ in English. Interestingly—and in keeping with the intuition about the geographical dispersion of treatment effects—changing schools reduces the adverse impact of commute distance on attendance by about half (0.093 days or 0.06 pp per mile). Whether a student is net better or worth off after a move depends on circumstances; the simple model here suggests a break-even commute distance of 8.8 miles, beyond which school changes are absence-minimizing. However, there does not seem to be an overall effect on tests scores.

Another potential channel through which the effects of distance may be transmitted is through distance's effect on homelessness outcomes. Housing instability and homelessness are associated with worse educational outcomes (though as the summary statistics on prior year school outcomes emphasize, homeless students' educational challenges predate shelter entry), so understanding the direct impact of distance on homelessness is important for policy. Hence, Table A.17 repeats Table 1 for homelessness outcomes.

The results are intuitive but surprising. Distant placements reduce shelter stays. The effect size is large—about 1.2–3.0 fewer days/mile for overall stays and 0.3–0.5 fewer days/per mile during the school year of shelter entry. At distances at the scale of changes-in-borough,

[9]Thus, there is one regression per column, which is different from the other result tables, which have one regression per cell.

distant placements reduce the lengths of overall shelter stays by about 22–43 days, or 5–9 percent of the average length of stay of 458 days. However, distant placements do not appear to change the likelihood of homelessness in subsequent school years. These results are consistent with the more detailed analysis of whole-family impacts of neighborhood-based shelter placements in Cassidy (2020).

Continuing this theme, Table A.18 studies the persistence of treatment effects. It repeats Table 1, Columns 1–3, for students in each of the two following school years (i.e, if t is the school year of shelter entry, these are years $t + 1$ and $t + 2$). There does seem to be some enduring impact of shelter assignment on school stability, at least in the short term. Students continue to be about 0.4 pp/mile more likely to change schools in year $t + 1$, and about 0.14 pp/mile more likely to change schools in year $t+2$, even as the baseline probability of changing schools diminishes by a third. On the other hand, shelter proximity does not appear to have much of a short-term legacy in terms attendance and test scores.

C References

Campbell, Kevin, Dennis Deck, and Carole Cox, Anton abd Broderick. 2018. "The Link King User Manual: Version v5.2."

Campbell, Kevin M. 2018. "The Link King: Record Linkage and Consolidation Software, v9.0."

Cassidy, Michael. 2020. "Short Moves and Long Stays: Homeless Family Responses to Exogenous Shelter Assignments in New York City." *IZA Discussion Papers*, No. 13559.

Freyaldenhoven, Simon, Christian Hansen, Jorge Pérez Pérez, and Jesse M Shapiro. 2021. "Visualization, Identification, and Estimation in the Linear Panel Event-study Design."

Google. 2023. "Google Maps Platform Distance Matrix API."

NYC Department of City Planning. 2017. "Geosupport System: User Programming Guide, Software Version 17.1."

NYC Department of Education. 2017. "2017 - 2018 School Locations." NYC Open Data.

Rambachan, Ashesh, and Jonathan Roth. 2023. "A More Credible Approach to Parallel Trends." *Review of Economic Studies*, rdad018.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's Trending in Difference-in-differences? A Synthesis of the Recent Econometrics Literature." *Journal of Econometrics*.

D Supplementary Tables

Table A.1: Panel Summary: Observations and School Years Per Student (Grades K–8)

Times Observed	Observations (Student-Years)			Students		
	(1) All	(2) Main: All	(3) Main: Sample	(4) All	(5) Main: All	(6) Main: Sample
1	70,631	26,660	26,660	6,930	127	24,179
2	63,701	26,533	2,481	9,480	1,069	2,289
3	54,221	25,464	192	9,141	1,926	173
4	45,080	23,538	19	8,572	2,572	18
5	36,508	20,966	1	7,532	2,885	1
6	28,976	18,081		7,082	3,503	
7	21,894	14,578		6,489	3,804	
8	15,405	10,774		6,192	4,096	
9	9,213	6,678		6,708	4,761	
10	2,505	1,917		2,106	1,603	
11	399	314		354	277	
12	45	37		45	37	
Total	348,578	175,540	29,353	70,631	26,660	26,660

Columns 1–3 give the number of student-school-years (which is the unit of observation) present in the data for students observed the row-delineated number of times. Columns 4–6 give the unique students observed the row-delineated number of times. Note that for observations, rows are cumulative in the sense that all being observed n times implies being observed $n-1$ times as well. However, for students, rows are mutually exclusive in the sense that students in row n are observed $> n-1$ but $< n+1$ times. "All" refers to the unrestricted dataset of student school years observed for students in grades K–8. "Main: All" refers to students in the main sample across the full set of school years observed, 2005–2016. "Main: Sample" refers only to student observations included in the main sample. Main sample consists school years of shelter entry among homeless public (non-charter) school K–8 students during school years 2010–2015 (who entered shelter during calendar years 2010–2016). To be included in the main sample, these students must also be enrolled in DOE 180 days prior to shelter entry and have non-missing attendance data; students whose shelter assignments are in the top five percentiles of commute distance from school (distance outliers) are also excluded.

Table A.2A: Summary Statistics by Housing Status, Grades K–8, 2010–2015

	Housing Status		
	Housed (1)	Homeless (2)	Difference (3)
Days Absent	11.216	27.287	16.071
	(12.399)	(20.868)	(0.100)
			[1.433]
Absence Rate	0.065	0.163	0.098
	(0.074)	(0.125)	(0.001)
			[1.517]
Changed School	0.108	0.385	0.277
	(0.310)	(0.486)	(0.002)
			[2.564]
Promoted	0.977	0.926	-0.051
	(0.151)	(0.263)	(0.001)
			[-0.052]
ELA Standardized Score (SD units)	0.010	-0.590	-0.600
	(0.998)	(0.956)	(0.006)
			[-57.507]
Math Standardized Score (SD units)	0.012	-0.675	-0.688
	(0.998)	(0.898)	(0.006)
			[-57.260]
Proficient	0.270	0.070	-0.200
	(0.444)	(0.256)	(0.001)
			[-0.740]
Manhattan	0.125	0.126	0.001
	(0.331)	(0.332)	(0.002)
			[0.009]
Bronx	0.215	0.434	0.219
	(0.411)	(0.496)	(0.003)
			[1.019]
Brooklyn	0.300	0.313	0.013
	(0.458)	(0.464)	(0.002)
			[0.042]
Queens	0.297	0.107	-0.189
	(0.457)	(0.310)	(0.002)
			[-0.638]
Staten Island	0.063	0.019	-0.044
	(0.243)	(0.138)	(0.001)
			[-0.693]
Grade	3.921	3.486	-0.435
	(2.585)	(2.567)	(0.012)
			[-0.111]
Obs.	3,907,657	82,400	

This table compares average characteristics of housed and homeless students. Rows list outcomes and covariates. Column one gives housed means. Column 2 gives homeless means. Column 3 gives the raw differences, estimated from separate bivariate regressions of each covariate on an indicator for homeless. Unit of observation is student-school-year. Students are classified as homeless during the school years in which they spend any time in homeless shelter; students not in shelter during a given school year are classified as housed. Sample consists of all NYC DOE K–8 grade students with non-missing admissions data during school years 2010–2015. Standard errors clustered at the family group level are given in parentheses. Percent changes relative to housed student means are given in brackets.

Table A.2B: Summary Statistics by Housing Status, Grades K–8, 2010–2015

	Housing Status		
	Housed (1)	Homeless (2)	Difference (3)
Black	0.244	0.529	0.285
	(0.430)	(0.499)	(0.003)
			[1.167]
Hispanic	0.413	0.427	0.014
	(0.492)	(0.495)	(0.003)
			[0.034]
White	0.163	0.024	-0.139
	(0.369)	(0.153)	(0.001)
			[-0.852]
Asian	0.167	0.008	-0.158
	(0.373)	(0.091)	(0.001)
			[-0.950]
Multi-Racial	0.006	0.004	-0.002
	(0.076)	(0.062)	(0.000)
			[-0.342]
Native American	0.008	0.008	0.000
	(0.087)	(0.089)	(0.000)
			[0.038]
Female	0.483	0.488	0.005
	(0.500)	(0.500)	(0.003)
			[0.010]
Student Age	9.623	9.422	-0.201
	(2.706)	(2.818)	(0.014)
			[-0.021]
IEP	0.192	0.282	0.090
	(0.394)	(0.450)	(0.002)
			[0.472]
District 75 (Disabilities)	0.020	0.034	0.015
	(0.139)	(0.182)	(0.001)
			[0.738]
ELL	0.165	0.102	-0.063
	(0.371)	(0.303)	(0.002)
			[-0.381]
Non-English	0.418	0.171	-0.247
	(0.493)	(0.377)	(0.002)
			[-0.591]
Foreign-Born	0.130	0.046	-0.083
	(0.336)	(0.211)	(0.001)
			[-0.642]
NYC-Born	0.799	0.779	-0.020
	(0.401)	(0.415)	(0.002)
			[-0.025]
Free or Reduced-Price Lunch	0.768	0.995	0.227
	(0.422)	(0.069)	(0.000)
			[0.296]
Obs.	3,907,657	82,400	

This table compares average characteristics of housed and homeless students. Rows list outcomes and covariates. Column one gives housed means. Column 2 gives homeless means. Column 3 gives the raw differences, estimated from separate bivariate regressions of each covariate on an indicator for homeless. Unit of observation is student-school-year. Students are classified as homeless during the school years in which they spend any time in homeless shelter; students not in shelter during a given school year are classified as housed. Sample consists of all NYC DOE K–8 grade students with non-missing attendance data during school years 2010–2015. Standard errors clustered at the family group level are given in parentheses. Percent changes relative to housed student means are given in brackets.

Table A.3A: Sample Means

	Main (1)	Student FE (2)	DID (3)
Days Absent	28.13	31.44	27.85
Absence Rate	0.166	0.187	0.160
Changed School	0.466	0.531	0.396
ELA Standardized Score (SD units)	-0.565	-0.659	-0.557
Math Standardized Score (SD units)	-0.661	-0.764	-0.645
Promoted	0.926	0.912	0.928
Commute Distance (miles)	8.240	8.117	8.400
Linear Distance (miles)	5.469	5.385	5.576
Out-of-School-Boro Treatment	0.450	0.448	0.458
Days Absent Prior Year	24.78	28.91	23.66
Absence Rate Prior Year	0.146	0.172	0.137
Changed School Prior Year	0.315	0.421	0.284
ELA Std. Score Prior Year	-0.555	-0.630	-0.542
Math Std. Score Prior Year	-0.660	-0.756	-0.643
Proficient Prior Year	0.103	0.067	0.102
Promoted Prior Year	0.917	0.898	0.916
Homeless Prior Year	0.265	0.463	0.008
Pre-Shelter School Distance	2.199	2.632	1.984
School Borough: Manhattan	0.117	0.123	0.111
School Borough: Bronx	0.402	0.432	0.401
School Borough: Brooklyn	0.339	0.341	0.349
School Borough: Queens	0.128	0.093	0.127
School Borough: Staten Island	0.014	0.010	0.013
Grade	3.806	3.798	3.926
Black	0.533	0.564	0.527
Hispanic	0.425	0.404	0.431
White	0.022	0.019	0.021
Asian	0.010	0.006	0.011
Multi-Racial	0.003	0.001	0.003
Native American	0.007	0.005	0.007
Female	0.493	0.501	0.492
Student Age	9.76	9.85	9.89
IEP	0.285	0.314	0.303
District 75 (Disabilities)	0.033	0.033	0.034
ELL	0.093	0.065	0.095
Non-English	0.169	0.124	0.177
Foreign-Born	0.047	0.024	0.049
NYC-Born	0.820	0.858	0.828
Youngest School-Age Child	0.472	0.405	0.470
Observations	29,353	5,150	16,252

Rows list outcomes, treatments, and covariates. Columns index samples of interest. Each cell gives the the mean for the row-indexed variable in the column-indexed sample. Column 1 is the main sample. Column 2 is the address fixed effects sample; specifically, it is the subsample of students with more than one homeless spell during the sample period, where the spells feature distinct school-shelter commute distances. Column 3 is the subsample of main sample students comprising the difference-in-differences sample.

Table A.3B: Sample Means

	Main (1)	Student FE (2)	DID (3)
Year	2012.51	2012.45	2012.83
Calendar Month of Shelter Entry	6.72	6.86	6.79
Female Head	0.923	0.927	0.921
Head Age	34.73	33.84	35.12
Head White	0.024	0.025	0.023
Head Black	0.528	0.557	0.522
Head Hispanic	0.407	0.397	0.413
Head Asian	0.005	0.003	0.005
Head Other Race	0.004	0.002	0.003
Head Unknown Race	0.033	0.017	0.035
Students in Family	2.330	2.417	2.346
Non-students in Family	2.088	2.149	2.073
Health Issue	0.337	0.311	0.371
Partner Present	0.268	0.286	0.266
Pregnant	0.044	0.036	0.042
Head Education: Less Than High School	0.593	0.629	0.596
Head Education: High School Grad	0.297	0.296	0.289
Head Education: Some College	0.048	0.042	0.048
Head Education: Unknown	0.061	0.033	0.067
On SNAP	0.732	0.777	0.748
Employed	0.390	0.383	0.401
On CA	0.362	0.400	0.362
Log Avg. Quarterly Earnings, Year Pre	2.774	2.625	2.881
Log Quarterly Earnings (|¿$0)	7.105	6.849	7.178
Eligibility: Eviction	0.462	0.411	0.508
Eligibility: Overcrowding	0.160	0.184	0.142
Eligibility: Conditions	0.066	0.061	0.061
Eligibility: DV	0.228	0.245	0.215
Eligibility: Other	0.083	0.099	0.073
Shelter Type: Tier II	0.551	0.539	0.558
Shelter Type: Commerical Hotel	0.174	0.165	0.169
Shelter Type: Family Cluster	0.270	0.288	0.270
Shelter Type: Other	0.005	0.009	0.003
School Enrollment	667.15	647.40	668.73
School Share Homeless	0.049	0.056	0.049
School Share Poor	0.879	0.888	0.867
School Share Black	0.418	0.429	0.414
School Share Hispanic	0.472	0.479	0.472
School Share White	0.042	0.034	0.044
School Share Asian	0.056	0.047	0.059
School Share Native American	0.009	0.009	0.010
School Share Multi-Racial	0.003	0.002	0.003
School Share Female	0.479	0.478	0.479
School Share ELL	0.153	0.154	0.154
School Share Asian	0.221	0.224	0.222
School Share Foreign-Born	0.116	0.110	0.120
School Share NYC-Born	0.806	0.810	0.803
Observations	29,353	5,150	16,252

Rows list outcomes, treatments, and covariates. Columns index samples of interest. Each cell gives the the mean for the row-indexed variable in the column-indexed sample. Column 1 is the main sample. Column 2 is the address fixed effects sample; specifically, it is the subsample of students with more than one homeless spell during the sample period, where the spells feature distinct school-shelter commute distances. Column 3 is the subsample of main sample students comprising the difference-in-differences sample.

Table A.4A: Description of Covariates

Variable	Description
A. Student and Family Covariates	
A1. Base Covariates	
School Year Fixed Effects	6 dummies for years 2010-11 to 2015-16.
Shelter-Entry-Month Fixed Effects	12 dummies for January–December.
School Borough Fixed Effects	5 dummies for school county origin (pre-shelter).
Grade Level Fixed Effects	9 dummies for grades K–8.
A2. Placement Covariates	
Family Students	Integer count of students in a family.
Family Non-Students	Integer count of non-student family members.
Partner Present	Dummy for head-of-household partner present in shelter.
Pregnant	Dummy for pregnant family member.
Health Issue	Dummy for family member with physical, mental, or substance abuse issue.
Student with Disability	Dummy for the presence of an Individualized Education Program (IEP).
District 75	Dummy for attending a D75 school. District 75 is a special non-geographical designation given to programs providing special services to students with significant disabilities.
Domestic Violence	Dummy for DV as reason for shelter eligibility.
Homeless Prior School Year*	Dummy for whether a student was homeless in the previous school year.
Pre-shelter School Distance*	Ventiles (20 groups) of linear distance between student's school and last residential address prior to shelter.
A3. Student Covariates	
Female	Dummy for female gender.
English Language Learner (ELL)	Dummy for English learner status.
Non-English Speaking	Dummy for non-English speaking home.
NYC Native	Dummy for NYC birthplace.
Foreign Born	Dummy for birthplace outside of the United States.
Youngest School-Aged Child	Dummy for whether a student is the youngest student in the family.
Age	Quadratic in monthly age at December 31 of a given school year.
Race Fixed Effects	Dummies for Black, Hispanic, White, Asian, Native American, multi-racial, and unknown.

This table describes all of the covariates included in the main analysis. Base covariates and placement covariates together comprise **"balance test" covariates**, conditional upon which shelter assignment is quasi-random. Collectively, student and family covariates (A) and student prior school year covariates (B) comprise **"main" covariates**, which define the baseline empirical specification. All levels of fixed effects and categorical variables are described for completeness, but base categories are omitted from regressions to avoid multicollinearity.

*All prior school year student covariates are appended with an additional "unknown" category to avoid dropping students with missing data.

+School covariates are calculated annually for each school using the 6.8 million student-school year observations comprising the complete DOE data (without imposing any sample restriction).

Table A.4B: Description of Covariates

Variable	Description
A4. Family Covariates	
Female	Dummy for female head-of-household.
Employed	Dummy for head employed during the year prior to shelter entry.
SNAP	Indicator for whether head had active Supplemental Nutrition Assistance Program case at shelter entry.
Cash Assistance	Indicator for whether head had active public assistance case at shelter entry.
Age	Quadratic in head age at time of shelter entry.
Log Average Quarterly Earnings	The natural logarithm of average quarterly earnings in the year prior to shelter entry, plus one dollar.
Educational Attainment Fixed Effects	Dummies for less than high school, high school graduate, some college or more, and unknown.
Shelter Type Fixed Effects	4 dummies for type of initial shelter assignment: Tier II, cluster unit, commercial hotel, and other.
Eligibility Fixed Effects	6 dummies for shelter eligibility reason: eviction, overcrowding, housing conditions, domestic violence, other, and unknown.
A5. School Covariates^{+}	
Enrollment	School-year-specific count of student enrollment.
Female Share	School-year-specific female student mean.
Black Share	School-year-specific Black student mean.
White Share	School-year-specific White student mean.
Hispanic Share	School-year-specific Hispanic student mean.
Asian Share	School-year-specific Asian student mean.
Native American Share	School-year-specific Native American student mean.
Multi-Race Share	School-year-specific multi-race student mean.
English Language Learner Share	School-year-specific English language learner student mean.
NYC-Native Share	School-year-specific NYC-native student mean.
Foreign-Born Share	School-year-specific foreign-born student mean.
Students with Disabilities Share	School-year-specific students with IEP mean.
Poor Share	School-year-specific mean of students receiving free or reduced-price lunch.
Homeless Share	School-year-specific homeless student mean.
B. Student Prior School Year Covariates*	
Days Absent	Ventiles of prior year days absent.
Absence	Ventiles of prior year absence rates.
School Change	Dummy for prior year school change.
School Change	Dummy for prior year proficiency in English and math.
Promoted	Indicator for promoted in the prior year.
C. School and Shelter Fixed Effects	
School Fixed Effects	1,148 dummies for school of origin.
Shelter Fixed Effects	226 dummies for initial shelter assignment.

This table describes all of the covariates included in the main analysis. Base covariates and placement covariates together comprise **"balance test" covariates**, conditional upon which shelter assignment is quasi-random. Collectively, student and family covariates (A) and student prior school year covariates (B) comprise **"main" covariates**, which define the baseline empirical specification. All levels of fixed effects and categorical variables are described for completeness, but base categories are omitted from regressions to avoid multicollinearity.

*All prior school year student covariates are appended with an additional "unknown" category to avoid dropping students with missing data.

$^{+}$School covariates are calculated annually for each school using the 6.8 million student-school year observations comprising the complete DOE data (without imposing any sample restriction).

Table A.5: Sample Step-Down Table

Refinement	Obs	Homeless Share
All Data	479,914	0.37
In-School (Grades K–8)	348,578	0.31
School Years 2010–2015	210,777	0.43
Excluding Charter Schools	194,405	0.43
Non-Missing Attendance	191,756	0.43
Enrolled in DOE Prior to Shelter	62,160	1.00
First School Year of Shelter Entry	31,886	1.00
Excluding Commute Distance Outliers	29,353	1.00

This table shows the path from the full data to the main analysis sample. Sample refinements are cumulative: each row imposes an additional restriction on the row above it. Data are from matched NYC Department of Homeless Services (calendar years 2010–2016) and Department of Education (school years 2005–2016) administrative records.

Table A.6: Match Stats: Students Age 5–18

	Students			Episodes		
Year of Birth	Obs	Matched	Match Rate	Obs	Matched	Match Rate
1992	493	341	0.69	499	343	0.69
1993	971	819	0.84	1,004	849	0.85
1994	1,577	1,390	0.88	1,720	1,518	0.88
1995	1,901	1,720	0.90	2,116	1,922	0.91
1996	2,390	2,179	0.91	2,778	2,539	0.91
1997	2,815	2,562	0.91	3,327	3,043	0.91
1998	3,501	3,202	0.91	4,219	3,875	0.92
1999	3,713	3,451	0.93	4,584	4,288	0.94
2000	4,022	3,676	0.91	4,886	4,493	0.92
2001	4,170	3,809	0.91	5,222	4,805	0.92
2002	4,246	3,875	0.91	5,292	4,879	0.92
2003	4,470	4,124	0.92	5,539	5,147	0.93
2004	4,938	4,523	0.92	6,216	5,753	0.93
2005	5,374	4,868	0.91	6,844	6,262	0.91
2006	5,544	5,017	0.90	7,020	6,425	0.92
2007	5,332	4,815	0.90	6,593	6,006	0.91
2008	5,287	4,735	0.90	6,366	5,757	0.90
2009	4,725	4,204	0.89	5,329	4,767	0.89
2010	4,062	3,576	0.88	4,380	3,876	0.88
2011	2,870	1,801	0.63	2,983	1,876	0.63
2012	1,657	41	0.02	1,665	42	0.03
Total	74,058	64,728	0.87	88,582	78,465	0.89

This table shows the results of probabilistic linkage of Department of Homeless Services (calendar year 2010–2016) and Department of Education (school year 2005–2016) administrative data. The sample universe is DHS family shelter entrants from calendar years 2010–2016. Candidate students are matched on first name, last name, date of birth (month and year), and sex. The statistics presented here encompass children ages 5–18 at some point during their homeless shelter spell.

Table A.7: Complete Sample Step-Down Table

Refinement	Obs	Homeless Share
All Data	6,798,801	0.02
In-School (Grades K–8)	4,396,886	0.02
School Years 2010–2015	4,396,886	0.02
Excluding Charter Schools	4,037,283	0.02
Non-Missing Attendance Outcomes	3,990,057	0.02

This table shows the path from the complete Department of Education data to the "complete" sample, which is analogous to the main sample but also includes housed students. Sample refinements are cumulative: each row imposes an additional restriction on the row above it. Data are from matched NYC Department of Homeless Services (calendar years 2010–2016) and Department of Education (school years 2005–2016) administrative records.

Table A.8A: Balance Test

	Out-of-School-Boro Treatment			Commute Distance	
	No	Yes	Raw Difference	Regression-Adjusted Diff.	
				Linear	Quad
	(1)	(2)	(3)	(4)	(5)
Commute Distance (miles)	3.917	13.528	9.611		
	(2.827)	(5.542)	(0.073)		
			{0.000}		
Days Absent Prior Year	24.771	24.801	0.030	-0.062	-0.236
	(18.790)	(19.159)	(0.299)	(0.234)	(0.270)
			{0.919}	{0.791}	{0.383}
			[0.001]	[-0.003]	[-0.010]
Absence Rate Prior Year	0.145	0.148	0.003	0.001	-0.001
	(0.111)	(0.117)	(0.002)	(0.001)	(0.002)
			{0.083}	{0.714}	{0.686}
			[0.022]	[0.004]	[-0.005]
Changed School Prior Year	0.308	0.325	0.017	-0.001	-0.007
	(0.462)	(0.468)	(0.007)	(0.005)	(0.006)
			{0.017}	{0.858}	{0.280}
			[0.054]	[-0.003]	[-0.022]
ELA Std. Score Prior Year	-0.568	-0.540	0.028	0.015	0.012
	(0.954)	(0.942)	(0.020)	(0.014)	(0.016)
			{0.166}	{0.267}	{0.448}
			[-0.050]	[-0.026]	[-0.022]
Math Std. Score Prior Year	-0.659	-0.661	-0.002	-0.006	0.002
	(0.896)	(0.976)	(0.020)	(0.014)	(0.017)
			{0.930}	{0.681}	{0.922}
			[0.003]	[0.009]	[-0.003]
Proficient Prior Year	0.108	0.097	-0.011	-0.002	0.000
	(0.310)	(0.296)	(0.005)	(0.004)	(0.005)
			{0.050}	{0.570}	{0.951}
			[-0.099]	[-0.022]	[0.003]
Promoted Prior Year	0.915	0.919	0.004	0.001	0.002
	(0.279)	(0.273)	(0.004)	(0.003)	(0.003)
			{0.207}	{0.705}	{0.629}
			[0.005]	[0.001]	[0.002]
Obs.	16,150	13,203			

This table provides a balance test of key student and family characteristics, all of which are covariates in the main analysis, across both borough and commute distance treatment definitions for the main sample. Rows list covariates. Column one gives in-borough means. Column 2 gives out-of-borough means. Column 3 gives the raw differences, estimated from separate bivariate regressions of each covariate on out-of-borough treatment. Columns 4 and 5 assess covariate balance by continuous commute distance (in miles), obtained by regressing each covariate separately on a linear (Column 4) or quadratic (Column 5) model in commute distance, controlling for balance test covariates (i.e., features conditional upon which shelter assignment is quasi-random). Reported contrasts compare covariates at the in-borough (3.9 miles) and out-of-borough (13.5 miles) commute distance means. Standard errors clustered at the family group level are given in parentheses. P-values for tests of null contrasts are given in braces. Percent changes relative to in-borough means are given in brackets. The first row summarizes commute distance means for students placed in shelters in and out of their school boroughs.

Table A.8B: Balance Test

	Out-of-School-Boro Treatment			Commute Distance	
	No	Yes	Raw Difference	Regression-Adjusted Diff.	
				Linear	Quad
	(1)	(2)	(3)	(4)	(5)
Female	0.495	0.491	-0.004	-0.003	-0.004
	(0.500)	(0.500)	(0.006)	(0.005)	(0.006)
			{0.501}	{0.531}	{0.476}
			[-0.008]	[-0.006]	[-0.008]
Student Age	9.775	9.741	-0.034	0.008	-0.000
	(2.686)	(2.699)	(0.033)	(0.007)	(0.008)
			{0.311}	{0.266}	{0.978}
			[-0.003]	[0.001]	[-0.000]
Black	0.529	0.537	0.007	-0.007	-0.023
	(0.499)	(0.499)	(0.008)	(0.006)	(0.007)
			{0.350}	{0.275}	{0.001}
			[0.014]	[-0.012]	[-0.044]
Hispanic	0.435	0.414	-0.022	-0.001	0.011
	(0.496)	(0.492)	(0.008)	(0.006)	(0.007)
			{0.006}	{0.812}	{0.106}
			[-0.049]	[-0.003]	[0.025]
White	0.019	0.025	0.006	0.005	0.007
	(0.137)	(0.157)	(0.002)	(0.002)	(0.002)
			{0.005}	{0.008}	{0.000}
			[0.306]	[0.236]	[0.384]
Non-English	0.168	0.171	0.003	0.005	0.010
	(0.373)	(0.376)	(0.006)	(0.005)	(0.005)
			{0.568}	{0.270}	{0.054}
			[0.020]	[0.030]	[0.062]
Foreign-Born	0.044	0.050	0.006	0.006	0.004
	(0.205)	(0.218)	(0.003)	(0.003)	(0.003)
			{0.054}	{0.035}	{0.172}
			[0.137]	[0.126]	[0.092]
NYC-Born	0.823	0.818	-0.005	-0.007	-0.003
	(0.382)	(0.386)	(0.006)	(0.005)	(0.005)
			{0.417}	{0.158}	{0.568}
			[-0.006]	[-0.008]	[-0.004]
Youngest School-Age Child	0.488	0.451	-0.037	-0.004	-0.004
	(0.500)	(0.498)	(0.006)	(0.004)	(0.005)
			{0.000}	{0.285}	{0.441}
			[-0.076]	[-0.009]	[-0.008]
Obs.	16,150	13,203			

This table provides a balance test of key student and family characteristics, all of which are covariates in the main analysis, across both borough and commute distance treatment definitions for the main sample. Rows list covariates. Column one gives in-borough means. Column 2 gives out-of-borough means. Column 3 gives the raw differences, estimated from separate bivariate regressions of each covariate on out-of-borough treatment. Columns 4 and 5 assess covariate balance by continuous commute distance (in miles), obtained by regressing each covariate separately on a linear (Column 4) or quadratic (Column 5) model in commute distance, controlling for balance test covariates (i.e., features conditional upon which shelter assignment is quasi-random). Reported contrasts compare covariates at the in-borough (3.9 miles) and out-of-borough (13.5 miles) commute distance means. Standard errors clustered at the family group level are given in parentheses. P-values for tests of null contrasts are given in braces. Percent changes relative to in-borough means are given in brackets.

Table A.8C: Balance Test

	Out-of-School-Boro Treatment			Commute Distance	
	No	Yes	Raw Difference	Regression-Adjusted Diff.	
				Linear	Quad
	(1)	(2)	(3)	(4)	(5)
Female Head	0.922	0.924	0.002	-0.001	0.002
	(0.268)	(0.264)	(0.004)	(0.004)	(0.004)
			{0.629}	{0.693}	{0.662}
			[0.002]	[-0.002]	[0.002]
Head Age	34.764	34.697	-0.067	0.166	0.181
	(7.480)	(7.258)	(0.112)	(0.083)	(0.093)
			{0.548}	{0.046}	{0.053}
			[-0.002]	[0.005]	[0.005]
On CA	0.366	0.356	-0.010	0.001	-0.004
	(0.482)	(0.479)	(0.008)	(0.006)	(0.007)
			{0.184}	{0.817}	{0.589}
			[-0.028]	[0.004]	[-0.010]
On SNAP	0.742	0.721	-0.021	0.001	-0.004
	(0.437)	(0.449)	(0.007)	(0.006)	(0.006)
			{0.002}	{0.899}	{0.559}
			[-0.029]	[0.001]	[-0.005]
Employed	0.396	0.383	-0.013	-0.009	-0.002
	(0.489)	(0.486)	(0.008)	(0.006)	(0.007)
			{0.089}	{0.120}	{0.746}
			[-0.033]	[-0.024]	[-0.006]
Log Quarterly Earnings (|¿$0)	7.109	7.100	-0.009	0.039	0.043
	(1.592)	(1.567)	(0.039)	(0.032)	(0.036)
			{0.820}	{0.224}	{0.234}
			[-0.001]	[0.005]	[0.006]
Head Education: Less Than High School	0.600	0.586	-0.014	-0.000	-0.006
	(0.490)	(0.493)	(0.008)	(0.006)	(0.007)
			{0.075}	{0.969}	{0.399}
			[-0.024]	[-0.000]	[-0.010]
Head Education: High School Grad	0.301	0.292	-0.009	-0.013	-0.011
	(0.459)	(0.455)	(0.007)	(0.006)	(0.007)
			{0.219}	{0.029}	{0.106}
			[-0.030]	[-0.043]	[-0.037]
Head Education: Some College	0.045	0.053	0.008	0.005	0.007
	(0.206)	(0.223)	(0.003)	(0.003)	(0.003)
			{0.021}	{0.086}	{0.033}
			[0.179]	[0.107]	[0.155]
Obs.	16,150	13,203			

This table provides a balance test of key student and family characteristics, all of which are covariates in the main analysis, across both borough and commute distance treatment definitions for the main sample. Rows list covariates. Column one gives in-borough means. Column 2 gives out-of-borough means. Column 3 gives the raw differences, estimated from separate bivariate regressions of each covariate on out-of-borough treatment. Columns 4 and 5 assess covariate balance by continuous commute distance (in miles), obtained by regressing each covariate separately on a linear (Column 4) or quadratic (Column 5) model in commute distance, controlling for balance test covariates (i.e., features conditional upon which shelter assignment is quasi-random). Reported contrasts compare covariates at the in-borough (3.9 miles) and out-of-borough (13.5 miles) commute distance means. Standard errors clustered at the family group level are given in parentheses. P-values for tests of null contrasts are given in braces. Percent changes relative to in-borough means are given in brackets.

Table A.9: Coefficient Stability: Covariate Specification Robustness

	None	Base	Balance Test	No Prior Year	Main	Shelter + School FE
	(1)	(2)	(3)	(4)	(5)	(6)
Days Absent	0.181	0.181	0.188	0.169	0.186	0.180
	(0.025)	(0.025)	(0.026)	(0.024)	(0.021)	(0.021)
	29,353	29,353	29,353	29,353	29,353	29,263
Absence Rate	0.0012	0.0012	0.0011	0.0011	0.0011	0.0011
	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
	29,353	29,353	29,353	29,353	29,353	29,263
Changed School	0.0150	0.0150	0.0114	0.0135	0.0115	0.0113
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
	29,353	29,353	29,353	29,353	29,353	29,263
ELA Standardized Score (SD units)	0.0011	0.0011	0.0009	0.0003	0.0008	0.0015
	(0.0012)	(0.0012)	(0.0012)	(0.0011)	(0.0011)	(0.0012)
	16,840	16,840	16,840	16,840	16,840	16,732
Math Standardized Score (SD units)	-0.0022	-0.0022	-0.0023	-0.0029	-0.0026	-0.0022
	(0.0012)	(0.0012)	(0.0012)	(0.0011)	(0.0011)	(0.0011)
	16,840	16,840	16,840	16,840	16,840	16,735
Promoted	0.0002	0.0002	0.0001	-0.0001	-0.0000	0.0000
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
	27,312	27,312	27,312	27,312	27,312	27,223
Base Covariates	No	Yes	Yes	Yes	Yes	Yes
Balance Test Covariates	No	No	Yes	Yes	Yes	Yes
Student and Family Covariates	No	No	No	Yes	Yes	Yes
Prior School Year Covariates	No	No	No	No	Yes	Yes
School and Shelter FE	No	No	No	No	No	Yes

This table assesses the sensitivity of Table 1 results to alternative covariate specifications. Outcomes are listed in rows. Analytical specifications are indexed by column. Treatment is continuous school-shelter commute distance (in miles). Unit of observation is a student school year. Each cell reports the coefficient on commute distance from a separate OLS linear regression of the row-indexed outcome on commute distance and the covariates described at the bottom of the table. Standard errors clustered by family group are given in parentheses. Sample sizes are given below standard errors.

Table A.10: Treatment Measure Robustness

	Commute Distance	Linear Distance	Commute Distance (Walk $\leq$ 0.5)	Transit Distance	Walk Distance	Commute Time
	(1)	(2)	(3)	(4)	(5)	(6)
Days Absent	0.186	0.278	0.186	0.186	0.228	0.05138
	(0.021)	(0.032)	(0.021)	(0.021)	(0.026)	(0.00585)
	29,353	29,353	29,353	29,337	29,352	27,877
Absence Rate	0.0011	0.0016	0.0011	0.0011	0.0013	0.00030
	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.00003)
	29,353	29,353	29,353	29,337	29,352	27,877
Changed School	0.0115	0.0169	0.0115	0.0115	0.0137	0.00315
	(0.0006)	(0.0008)	(0.0006)	(0.0006)	(0.0007)	(0.00016)
	29,353	29,353	29,353	29,337	29,352	27,877
ELA Standardized Score (SD units)	0.0008	0.0010	0.0008	0.0008	0.0012	0.00004
	(0.0011)	(0.0016)	(0.0011)	(0.0011)	(0.0013)	(0.00031)
	16,840	16,840	16,840	16,830	16,839	16,026
Math Standardized Score (SD units)	-0.0026	-0.0040	-0.0026	-0.0025	-0.0030	-0.00059
	(0.0011)	(0.0016)	(0.0011)	(0.0011)	(0.0013)	(0.00031)
	16,840	16,840	16,840	16,831	16,839	16,031
Promoted	-0.0000	-0.0001	-0.0000	0.0000	-0.0001	-0.00006
	(0.0003)	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.00008)
	27,312	27,312	27,312	27,296	27,311	25,934
Treatment Mean (SD)	8.24	5.47	8.25	8.25	6.63	46.34
	(6.41)	(4.22)	(6.40)	(6.40)	(5.21)	(22.56)
Units	Miles	Miles	Miles	Miles	Miles	Minutes
Student and Family Covariates	Yes	Yes	Yes	Yes	Yes	Yes
School and Shelter FE	No	No	No	No	No	No

This table assesses the robustness of Table 1 results to alternative treatment measures . Outcomes are listed in rows. Treatment measures are indexed by column. Unit of observation is a student school year. Each cell reports the coefficient on treatment from a separate OLS linear regression of the row-indexed outcome on treatment and the main covariate specification from Table 1. Column 1 repeats Column 2 from Table 1 for reference, using the main treatment definition of continuous school-shelter commute distance (in miles). Column 2 defines treatment as linear (Cartesian) distance (in miles). Column 3 uses commute distance, but changes the walk threshold to $\leq$ 0.5 miles. Column 4 uses public transit distance (in miles). Column 5 uses walking distance (in miles). Column 6 uses commute time (in minutes). Standard errors clustered by family group are given in parentheses. Sample sizes are given below standard errors. The last set of numerical rows gives the treatment mean and standard deviation.

Table A.11: Family Selection and Alternative Samples

	Main	Remain in Initial Shelter	LOS $\geq$ 30 Days	First Spell Only	Including Outliers	All Stayers
	(1)	(2)	(3)	(4)	(5)	(6)
Days Absent	0.186	0.199	0.189	0.188	0.198	0.203
	(0.021)	(0.023)	(0.021)	(0.023)	(0.020)	(0.025)
	29,353	23,372	27,880	22,871	30,898	18,092
Absence Rate	0.0011	0.0012	0.0011	0.0011	0.0010	0.0011
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
	29,353	23,372	27,880	22,871	30,898	18,092
Changed School	0.0115	0.0130	0.0120	0.0119	0.0093	0.0116
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0005)	(0.0007)
	29,353	23,372	27,880	22,871	30,898	18,092
ELA Standardized Score (SD units)	0.0008	0.0002	0.0011	0.0013	0.0005	0.0007
	(0.0011)	(0.0012)	(0.0011)	(0.0012)	(0.0010)	(0.0014)
	16,840	13,409	16,071	13,442	17,707	10,661
Math Standardized Score (SD units)	-0.0026	-0.0028	-0.0024	-0.0019	-0.0024	-0.0017
	(0.0011)	(0.0012)	(0.0011)	(0.0012)	(0.0010)	(0.0014)
	16,840	13,402	16,080	13,452	17,709	10,676
Promoted	-0.0000	-0.0001	0.0001	-0.0001	-0.0000	-0.0001
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)
	27,312	21,629	25,984	21,358	28,726	16,798
Student and Family Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Prior School Year Covariates	Yes	Yes	Yes	Yes	Yes	Yes
School and Shelter FE	No	No	No	No	No	No

This table addresses potential concerns about family self-selection into shelter assignments by assessing the robustness of Table 1 results to alternative samples. Outcomes are listed in rows. Samples are indexed by column. Unit of observation is a student school year. Each cell reports the coefficient on continuous commute distance (in miles) from a separate OLS linear regression of the row-indexed outcome on commute distance and the main covariate specification from Table 1. Column 1 repeats them main sample results from Column 2 of Table 1 for reference. Column 2 limits the main sample to students whose primary shelter (i.e., shelter of longest stay) during their shelter spell is their intial shelter, thus excluding students from families with major shelter changes. Column 3 limits the main sample to students whose shelter stays are at least 30 days, thus excluding students from families with short stays. Column 4 limits the main sample to students' first observed homeless spell (post-2010), thus excluding subsequent returns to shelter. Column 5 expands the main sample to include commute distance outliers—that is, it includes students placed in the 95th+ percentiles. Column 6 includes all of the preceding sample modifications at once. Standard errors clustered by family group are given in parentheses. Sample sizes are given below standard errors.

Table A.12: Alternative Proficiency Measures

		Commute Distance					Out-of-Borough				
	Mean	Base	Main	S+S FE	Stud. FE	Full	Base	Main	S+S FE	Stud. FE	Full
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Math Proficient	0.1583	-0.0008	-0.0008	-0.0007	0.0008	0.0044	-0.0125	-0.0112	-0.0124	-0.0068	0.0290
	(0.3650)	(0.0004)	(0.0004)	(0.0004)	(0.0012)	(0.0021)	(0.0056)	(0.0053)	(0.0056)	(0.0173)	(0.0305)
	18,650	18,650	18,650	18,546	2,636	2,226	18,650	18,650	18,546	2,636	2,226
ELA Proficient	0.1374	-0.0007	-0.0006	-0.0005	0.0009	-0.0002	-0.0062	-0.0055	-0.0051	-0.0008	-0.0225
	(0.3443)	(0.0004)	(0.0004)	(0.0004)	(0.0013)	(0.0018)	(0.0054)	(0.0051)	(0.0054)	(0.0171)	(0.0248)
	18,650	18,650	18,650	18,546	2,636	2,226	18,650	18,650	18,546	2,636	2,226
Proficient	0.0809	-0.0009	-0.0008	-0.0007	0.0006	0.0013	-0.0095	-0.0085	-0.0083	-0.0037	0.0122
	(0.2727)	(0.0003)	(0.0003)	(0.0003)	(0.0010)	(0.0016)	(0.0043)	(0.0040)	(0.0042)	(0.0133)	(0.0220)
	18,650	18,650	18,650	18,546	2,636	2,226	18,650	18,650	18,546	2,636	2,226
ELA Proficient (Test Takers Only)	0.1522	-0.0007	-0.0008	-0.0005	0.0007	0.0012	-0.0062	-0.0062	-0.0056	0.0012	-0.0024
	(0.3592)	(0.0005)	(0.0004)	(0.0005)	(0.0015)	(0.0025)	(0.0059)	(0.0056)	(0.0060)	(0.0192)	(0.0322)
	16,840	16,840	16,840	16,732	2,169	1,781	16,840	16,840	16,732	2,169	1,781
Math Proficient (Test Takers Only)	0.1753	-0.0008	-0.0008	-0.0007	0.0013	0.0073	-0.0129	-0.0126	-0.0137	0.0034	0.0455
	(0.3802)	(0.0005)	(0.0004)	(0.0005)	(0.0014)	(0.0027)	(0.0061)	(0.0059)	(0.0062)	(0.0195)	(0.0381)
	16,840	16,840	16,840	16,735	2,158	1,755	16,840	16,840	16,735	2,158	1,755
Proficient (Test Takers Only)	0.0913	-0.0009	-0.0009	-0.0007	0.0007	0.0042	-0.0102	-0.0097	-0.0092	-0.0059	0.0346
	(0.2881)	(0.0004)	(0.0003)	(0.0004)	(0.0012)	(0.0024)	(0.0048)	(0.0045)	(0.0048)	(0.0161)	(0.0310)
	16,522	16,522	16,522	16,414	2,087	1,698	16,522	16,522	16,414	2,087	1,698
ELA Scale Score	405.1690	0.0157	0.0340	0.0610	-0.0172	-0.0740	0.1286	0.3659	0.2756	0.7673	-1.2436
	(179.8327)	(0.0354)	(0.0346)	(0.0376)	(0.0899)	(0.1849)	(0.4601)	(0.4487)	(0.4833)	(1.1717)	(2.4459)
	16,840	16,840	16,840	16,732	2,169	1,781	16,840	16,840	16,732	2,169	1,781
Math Scale Score	407.7378	-0.1064	-0.0933	-0.0806	-0.1957	-0.3733	-1.5834	-1.4449	-1.3829	-2.4264	-4.7838
	(187.0625)	(0.0412)	(0.0397)	(0.0424)	(0.1087)	(0.1840)	(0.5276)	(0.5067)	(0.5440)	(1.4421)	(2.7535)
	16,840	16,840	16,840	16,735	2,158	1,755	16,840	16,840	16,735	2,158	1,755

This table assess the sensitivity of the main results about student performance on grade 3–8 state tests in English and math to alternative measures of proficiency. Outcomes are listed in rows. Unit of observation is a student school year. Repeating the layout of Table 1, each cell reports the coefficient on treatment from a separate OLS linear regression of the row-indexed outcome on treatment and the covariate specifications described at the bottom of the table. Rows 1–3 use a binary measure that defines proficient as scoring level 3 or 4 on state tests and not proficient as scoring levels 1 or 2 or missing the test. Rows 4–6 use the same binary defintion of proficiency, but exclude students who missed tests. Rows 7 and 8 use numerical scale scores (in points) as the outcomes. Standard errors clustered by family group are given in parentheses. Sample sizes are given below standard errors.

Table A.13: Difference-in-Differences: Additional Results

	Out-of-Borough Treatment			Linear Distance Change		
	OLS	TWFE	TWFE-Pre	OLS	TWFE	TWFE-Pre
	(1)	(2)	(3)	(4)	(5)	(6)
Days Absent	2.286	2.305	3.028	0.273	0.299	0.324
	(0.321)	(0.317)	(0.511)	(0.038)	(0.034)	(0.056)
	16,198	32,054	11,198	16,198	32,054	11,198
Absence Rate	0.0146	0.0131	0.0167	0.0017	0.0018	0.0019
	(0.0018)	(0.0019)	(0.0029)	(0.0002)	(0.0002)	(0.0003)
	16,198	32,052	11,196	16,198	32,052	11,196
Changed School	0.1424	0.1606	0.1384	0.0165	0.0133	0.0102
	(0.0096)	(0.0120)	(0.0183)	(0.0011)	(0.0013)	(0.0020)
	16,198	32,396	11,200	16,198	32,396	11,200
ELA Standardized Score (SD units)	0.0182	0.0001	-0.0073	0.0017	0.0019	0.0022
	(0.0171)	(0.0154)	(0.0212)	(0.0020)	(0.0016)	(0.0022)
	9,961	15,414	7,624	9,961	15,414	7,624
Math Standardized Score (SD units)	-0.0430	-0.0053	-0.0302	-0.0030	0.0014	0.0017
	(0.0173)	(0.0174)	(0.0227)	(0.0020)	(0.0019)	(0.0024)
	9,909	15,502	7,512	9,909	15,502	7,512
Promoted	0.0019	0.0021	-0.0045	0.0001	0.0001	-0.0005
	(0.0046)	(0.0044)	(0.0056)	(0.0005)	(0.0005)	(0.0006)
	15,254	30,506	10,620	15,254	30,506	10,620
Sample	DID-Main	DID	DID-Pre	DID-Main	DID	DID-Pre
Covariates	Main	YBG	YBG	Main	YBG	YBG

Outcomes are given in rows. Estimation methods are indexed by column. Unit of observation is a student school year. Each cell reports the average treatment effect on the treated (ATT) of the supercolumn-indexed treatment on the row-indexed outcome from a separate estimation using the column-indexed method. The sample and covariates for each method are summarized at the bottom of the table. Columns 1 and 4 repeat the main OLS specification from Table 1 for the main sample subsample of the DID sample (i.e., shelter entry years only), corresponding to Table 1, Columns 6 and 2, respectively. Columns 2 and 4 give the standard two-way fixed effects DID estimates, controlling for student and relative time fixed effects. Columns 3 and 6 limit the DID sample to the students observed continuously for relative-time school years, $-2 \leq R \leq 1$, where $R = 1$ in the treated school year and estimate the TWFE model for this sub-subsample; these are the students who are included in the pre-trends analysis. Each cell in Columns 1–3 reports results for treatment defined as an indicator for out-of-school-borough shelter placement. Each cell in Columns 4–6 reports the coefficient on continuous linear school-shelter commute distance (in miles). Standard errors clustered by family group are given in parentheses. Sample sizes are given below standard errors.

Table A.14A: Heterogeneity Analysis

	Days Absent (1)	Absence Rate (2)	Changed School (3)	ELA Standardized (4)	Math Standardized (5)	Promoted (6)
Commute Distance Quantile						
Q1	0.423	0.0021	0.0259	-0.0058	-0.0157	-0.0090
	(0.295)	(0.0017)	(0.0084)	(0.0166)	(0.0176)	(0.0045)
	7,342	7,342	7,342	4,281	4,301	6,888
Q2	0.295	0.0020	0.0115	0.0030	-0.0337	-0.0014
	(0.248)	(0.0015)	(0.0073)	(0.0156)	(0.0154)	(0.0039)
	7,336	7,336	7,336	4,231	4,210	6,793
Q3	0.142	0.0009	0.0127	-0.0044	-0.0153	-0.0016
	(0.096)	(0.0006)	(0.0027)	(0.0053)	(0.0053)	(0.0014)
	7,338	7,338	7,338	4,181	4,193	6,825
Q4	0.019	0.0005	0.0102	-0.0008	0.0053	-0.0006
	(0.105)	(0.0006)	(0.0029)	(0.0057)	(0.0058)	(0.0015)
	7,337	7,337	7,337	4,147	4,136	6,806
Difference in Means	-0.404	-0.0016	-0.0157	0.0050	0.0210	0.0084
	(0.313)	(0.0018)	(0.0089)	(0.0176)	(0.0185)	(0.0047)
	0.1969	0.3853	0.0779	0.7748	0.2571	0.0753
Days Absent Prior Year Quartile						
Q1	0.152	0.0009	0.0106	0.0000	-0.0018	-0.0002
	(0.029)	(0.0002)	(0.0011)	(0.0019)	(0.0020)	(0.0004)
	6,805	6,805	6,805	4,665	4,649	6,426
Q2	0.231	0.0014	0.0116	-0.0009	-0.0055	-0.0001
	(0.033)	(0.0002)	(0.0011)	(0.0020)	(0.0021)	(0.0006)
	6,760	6,760	6,760	4,171	4,156	6,360
Q3	0.199	0.0012	0.0141	-0.0003	-0.0025	0.0003
	(0.040)	(0.0002)	(0.0011)	(0.0023)	(0.0023)	(0.0006)
	6,251	6,251	6,251	3,583	3,559	5,885
Q4	0.173	0.0011	0.0130	0.0052	-0.0021	-0.0003
	(0.054)	(0.0003)	(0.0011)	(0.0027)	(0.0026)	(0.0007)
	6,301	6,301	6,301	3,077	3,049	5,866
Unknown	0.178	0.0008	0.0048	0.0007	-0.0006	0.0005
	(0.061)	(0.0004)	(0.0010)	(0.0042)	(0.0040)	(0.0009)
	3,236	3,236	3,236	1,344	1,427	2,775
Difference in Means	0.021	0.0002	0.0023	0.0051	-0.0003	-0.0001
	(0.061)	(0.0003)	(0.0016)	(0.0033)	(0.0033)	(0.0008)
	0.7333	0.5641	0.1399	0.1215	0.9368	0.8710
Changed School Prior Year						
Yes	0.171	0.0011	0.0136	0.0003	-0.0029	0.0003
	(0.037)	(0.0002)	(0.0010)	(0.0022)	(0.0021)	(0.0005)
	8,494	8,494	8,494	4,489	4,478	7,839
No	0.196	0.0011	0.0115	0.0010	-0.0026	-0.0001
	(0.024)	(0.0001)	(0.0007)	(0.0013)	(0.0013)	(0.0003)
	18,437	18,437	18,437	11,209	11,143	17,411
Difference in Means	-0.025	-0.0001	0.0021	-0.0007	-0.0003	0.0005
	(0.045)	(0.0003)	(0.0012)	(0.0025)	(0.0025)	(0.0006)
	0.5765	0.8425	0.0805	0.7926	0.8949	0.4632

This table conducts a heterogeneity analysis by repeating the main analysis from Table 1 for subgroups. Unit of observation is a student school year. Outcomes are listed in columns. Rows index the characteristics and levels defining the subsamples among which the heterogeneity analysis is conducted. Each cell in a characteristic-level row reports the coefficient on continuous commute distance (in miles) from a separate regression of the column-enumerated outcome on commute distance and the main covariate specification from Table 1 for the subsample defined by the characteristic-level row. Standard errors clustered by family group are given in parentheses. Sample sizes are given in the third row of each cell. Difference in Means row for each characteristic gives the difference in coefficients, standard errors of the differences (in parentheses), and p-values (in the third row). For binary characteristics, the comparison is between the two levels; for ordered categorical variables, the comparison is between the highest and lowest levels.

Table A.14B: Heterogeneity Analysis

	Days Absent (1)	Absence Rate (2)	Changed School (3)	ELA Standardized (4)	Math Standardized (5)	Promoted (6)
Youngest School-Age Child						
Yes	0.191	0.0011	0.0121	0.0020	-0.0020	0.0003
	(0.025)	(0.0001)	(0.0007)	(0.0018)	(0.0017)	(0.0004)
	13,841	13,841	13,841	6,792	6,772	12,777
No	0.184	0.0011	0.0109	0.0002	-0.0030	-0.0002
	(0.028)	(0.0002)	(0.0007)	(0.0014)	(0.0014)	(0.0004)
	15,512	15,512	15,512	10,048	10,068	14,535
Difference in Means	0.007	-0.0000	0.0012	0.0018	0.0010	0.0006
	(0.037)	(0.0002)	(0.0010)	(0.0022)	(0.0022)	(0.0005)
	0.8419	0.9705	0.2134	0.4113	0.6398	0.3053
1 Student in Family						
Yes	0.185	0.0011	0.0137	0.0020	-0.0026	0.0002
	(0.034)	(0.0002)	(0.0009)	(0.0023)	(0.0024)	(0.0005)
	8,496	8,496	8,496	4,136	4,118	7,907
No	0.190	0.0011	0.0107	0.0005	-0.0026	-0.0002
	(0.025)	(0.0001)	(0.0007)	(0.0012)	(0.0012)	(0.0003)
	20,857	20,857	20,857	12,704	12,722	19,405
Difference in Means	-0.004	-0.0000	0.0030	0.0015	0.0000	0.0004
	(0.043)	(0.0003)	(0.0011)	(0.0026)	(0.0027)	(0.0006)
	0.9205	0.9320	0.0066	0.5814	0.9941	0.5089
IEP						
Yes	0.214	0.0013	0.0114	0.0023	-0.0029	0.0004
	(0.037)	(0.0002)	(0.0009)	(0.0023)	(0.0022)	(0.0004)
	8,358	8,358	8,358	4,973	4,947	7,797
No	0.171	0.0010	0.0115	0.0001	-0.0024	-0.0001
	(0.023)	(0.0001)	(0.0006)	(0.0011)	(0.0012)	(0.0003)
	20,995	20,995	20,995	11,867	11,893	19,515
Difference in Means	0.043	0.0003	-0.0001	0.0022	-0.0005	0.0005
	(0.043)	(0.0003)	(0.0011)	(0.0026)	(0.0025)	(0.0005)
	0.3192	0.2030	0.9532	0.3875	0.8298	0.3663
NYC-Born						
Yes	0.182	0.0011	0.0117	0.0003	-0.0025	-0.0000
	(0.023)	(0.0001)	(0.0006)	(0.0012)	(0.0012)	(0.0003)
	24,083	24,083	24,083	13,640	13,561	22,675
No	0.207	0.0012	0.0107	0.0030	-0.0033	0.0001
	(0.043)	(0.0003)	(0.0012)	(0.0026)	(0.0025)	(0.0006)
	5,270	5,270	5,270	3,200	3,279	4,637
Difference in Means	-0.025	-0.0002	0.0010	-0.0027	0.0008	-0.0001
	(0.049)	(0.0003)	(0.0014)	(0.0028)	(0.0028)	(0.0007)
	0.6028	0.5325	0.4514	0.3412	0.7816	0.8564

This table conducts a heterogeneity analysis by repeating the main analysis from Table 1 for subgroups. Unit of observation is a student school year. Outcomes are listed in columns. Rows index the characteristics and levels defining the subsamples among which the heterogeneity analysis is conducted. Each cell in a characteristic-level row reports the coefficient on continuous commute distance (in miles) from a separate regression of the column-enumerated outcome on commute distance and the main covariate specification from Table 1 for the subsample defined by the characteristic-level row. Standard errors clustered by family group are given in parentheses. Sample sizes are given in the third row of each cell. Difference in Means row for each characteristic gives the difference in coefficients, standard errors of the differences (in parentheses), and p-values (in the third row). For binary characteristics, the comparison is between the two levels; for ordered categorical variables, the comparison is between the highest and lowest levels.

Table A.14C: Heterogeneity Analysis

	Days Absent (1)	Absence Rate (2)	Changed School (3)	ELA Standardized (4)	Math Standardized (5)	Promoted (6)
Female						
Yes	0.184	0.0011	0.0122	0.0003	-0.0014	-0.0002
	(0.026)	(0.0002)	(0.0007)	(0.0014)	(0.0014)	(0.0004)
	14,476	14,476	14,476	8,494	8,487	13,514
No	0.186	0.0011	0.0109	0.0012	-0.0038	0.0002
	(0.027)	(0.0002)	(0.0007)	(0.0016)	(0.0016)	(0.0004)
	14,877	14,877	14,877	8,346	8,353	13,798
Difference in Means	-0.002	0.0000	0.0013	-0.0009	0.0023	-0.0004
	(0.038)	(0.0002)	(0.0010)	(0.0021)	(0.0021)	(0.0005)
	0.9666	0.8887	0.2118	0.6675	0.2811	0.4961
Black						
Yes	0.185	0.0011	0.0120	0.0010	-0.0024	-0.0003
	(0.026)	(0.0002)	(0.0007)	(0.0015)	(0.0014)	(0.0004)
	15,631	15,631	15,631	9,136	9,081	14,618
No	0.187	0.0011	0.0111	0.0005	-0.0031	0.0004
	(0.032)	(0.0002)	(0.0008)	(0.0016)	(0.0017)	(0.0004)
	13,722	13,722	13,722	7,704	7,759	12,694
Difference in Means	-0.002	-0.0000	0.0009	0.0005	0.0007	-0.0008
	(0.041)	(0.0002)	(0.0011)	(0.0022)	(0.0022)	(0.0006)
	0.9708	0.8756	0.4374	0.8257	0.7420	0.1726
Hispanic						
Yes	0.191	0.0012	0.0114	0.0007	-0.0037	0.0004
	(0.032)	(0.0002)	(0.0009)	(0.0017)	(0.0017)	(0.0005)
	12,486	12,486	12,486	7,049	7,106	11,556
No	0.183	0.0010	0.0117	0.0004	-0.0021	-0.0003
	(0.026)	(0.0002)	(0.0007)	(0.0014)	(0.0014)	(0.0003)
	16,867	16,867	16,867	9,791	9,734	15,756
Difference in Means	0.008	0.0001	-0.0003	0.0003	-0.0016	0.0007
	(0.042)	(0.0002)	(0.0011)	(0.0022)	(0.0022)	(0.0006)
	0.8394	0.6097	0.7882	0.8877	0.4700	0.2557
White						
Yes	0.045	-0.0001	0.0086	-0.0128	0.0021	-0.0007
	(0.149)	(0.0009)	(0.0037)	(0.0098)	(0.0092)	(0.0020)
	643	643	643	346	349	590
No	0.187	0.0011	0.0116	0.0009	-0.0025	0.0000
	(0.021)	(0.0001)	(0.0006)	(0.0011)	(0.0011)	(0.0003)
	28,710	28,710	28,710	16,494	16,491	26,722
Difference in Means	-0.141	-0.0012	-0.0030	-0.0137	0.0047	-0.0007
	(0.150)	(0.0009)	(0.0037)	(0.0098)	(0.0092)	(0.0020)
	0.3470	0.1711	0.4200	0.1635	0.6121	0.7329

This table conducts a heterogeneity analysis by repeating the main analysis from Table 1 for subgroups. Unit of observation is a student school year. Outcomes are listed in columns. Rows index the characteristics and levels defining the subsamples among which the heterogeneity analysis is conducted. Each cell in a characteristic-level row reports the coefficient on continuous commute distance (in miles) from a separate regression of the column-enumerated outcome on commute distance and the main covariate specification from Table 1 for the subsample defined by the characteristic-level row. Standard errors clustered by family group are given in parentheses. Sample sizes are given in the third row of each cell. Difference in Means row for each characteristic gives the difference in coefficients, standard errors of the differences (in parentheses), and p-values (in the third row). For binary characteristics, the comparison is between the two levels; for ordered categorical variables, the comparison is between the highest and lowest levels.

Table A.14D: Heterogeneity Analysis

	Days Absent (1)	Absence Rate (2)	Changed School (3)	ELA Standardized (4)	Math Standardized (5)	Promoted (6)
Head Education: Less Than High School						
Yes	0.189	0.0011	0.0117	0.0014	-0.0028	-0.0002
	(0.027)	(0.0002)	(0.0007)	(0.0014)	(0.0014)	(0.0004)
	17,417	17,417	17,417	9,906	9,914	16,239
No	0.182	0.0010	0.0112	0.0002	-0.0017	0.0003
	(0.031)	(0.0002)	(0.0009)	(0.0016)	(0.0016)	(0.0004)
	11,936	11,936	11,936	6,934	6,926	11,073
Difference in Means	0.008	0.0001	0.0005	0.0011	-0.0011	-0.0005
	(0.042)	(0.0002)	(0.0011)	(0.0022)	(0.0022)	(0.0005)
	0.8562	0.5471	0.6816	0.5953	0.6199	0.3845
Head Education: High School Grad						
Yes	0.186	0.0010	0.0114	0.0006	-0.0019	-0.0000
	(0.038)	(0.0002)	(0.0010)	(0.0019)	(0.0019)	(0.0005)
	8,728	8,728	8,728	5,094	5,084	8,065
No	0.187	0.0011	0.0115	0.0009	-0.0027	0.0000
	(0.025)	(0.0001)	(0.0007)	(0.0013)	(0.0013)	(0.0003)
	20,625	20,625	20,625	11,746	11,756	19,247
Difference in Means	-0.001	-0.0002	-0.0001	-0.0003	0.0007	-0.0001
	(0.045)	(0.0003)	(0.0012)	(0.0023)	(0.0023)	(0.0006)
	0.9875	0.5173	0.9041	0.8924	0.7551	0.8860
Head Education: Some College						
Yes	0.178	0.0012	0.0072	0.0036	-0.0007	0.0010
	(0.081)	(0.0005)	(0.0024)	(0.0052)	(0.0048)	(0.0012)
	1,414	1,414	1,414	830	821	1,324
No	0.185	0.0011	0.0117	0.0005	-0.0028	-0.0001
	(0.021)	(0.0001)	(0.0006)	(0.0011)	(0.0011)	(0.0003)
	27,939	27,939	27,939	16,010	16,019	25,988
Difference in Means	-0.007	0.0001	-0.0044	0.0031	0.0020	0.0011
	(0.084)	(0.0005)	(0.0025)	(0.0053)	(0.0049)	(0.0012)
	0.9311	0.8604	0.0765	0.5602	0.6834	0.3539
Employed						
Yes	0.191	0.0011	0.0113	0.0021	-0.0019	-0.0001
	(0.030)	(0.0002)	(0.0009)	(0.0017)	(0.0017)	(0.0004)
	11,453	11,453	11,453	6,589	6,564	10,700
No	0.182	0.0011	0.0117	-0.0001	-0.0033	0.0001
	(0.028)	(0.0002)	(0.0007)	(0.0014)	(0.0014)	(0.0004)
	17,900	17,900	17,900	10,251	10,276	16,612
Difference in Means	0.009	0.0000	-0.0004	0.0021	0.0014	-0.0002
	(0.041)	(0.0002)	(0.0011)	(0.0022)	(0.0022)	(0.0006)
	0.8301	0.9813	0.7520	0.3310	0.5234	0.6907
Health Issue						
Yes	0.208	0.0011	0.0118	0.0022	-0.0027	0.0002
	(0.037)	(0.0002)	(0.0009)	(0.0018)	(0.0018)	(0.0005)
	9,891	9,891	9,891	5,755	5,740	9,226
No	0.174	0.0011	0.0112	0.0001	-0.0024	-0.0001
	(0.025)	(0.0001)	(0.0007)	(0.0013)	(0.0014)	(0.0003)
	19,462	19,462	19,462	11,085	11,100	18,086
Difference in Means	0.034	0.0001	0.0006	0.0020	-0.0003	0.0003
	(0.044)	(0.0003)	(0.0012)	(0.0023)	(0.0022)	(0.0006)
	0.4381	0.8070	0.6044	0.3727	0.8773	0.6301

This table conducts a heterogeneity analysis by repeating the main analysis from Table 1 for subgroups. Unit of observation is a student school year. Outcomes are listed in columns. Rows index the characteristics and levels defining the subsamples among which the heterogeneity analysis is conducted. Each cell in a characteristic-level row reports the coefficient on continuous commute distance (in miles) from a separate regression of the column-enumerated outcome on commute distance and the main covariate specification from Table 1 for the subsample defined by the characteristic-level row. Standard errors clustered by family group are given in parentheses. Sample sizes are given in the third row of each cell. Difference in Means row for each characteristic gives the difference in coefficients, standard errors of the differences (in parentheses), and p-values (in the third row). For binary characteristics, the comparison is between the two levels; for ordered categorical variables, the comparison is between the highest and lowest levels.

Table A.14E: Heterogeneity Analysis

	Days Absent (1)	Absence Rate (2)	Changed School (3)	ELA Standardized (4)	Math Standardized (5)	Promoted (6)
School Mean Days Absent Above Median						
Yes	0.214	0.0012	0.0114	0.0005	-0.0029	-0.0003
	(0.026)	(0.0002)	(0.0007)	(0.0013)	(0.0013)	(0.0003)
	18,947	18,947	18,947	10,874	10,862	17,636
No	0.137	0.0008	0.0121	0.0009	-0.0026	0.0005
	(0.032)	(0.0002)	(0.0009)	(0.0018)	(0.0018)	(0.0005)
	10,406	10,406	10,406	5,966	5,978	9,676
Difference in Means	0.076	0.0004	-0.0007	-0.0004	-0.0003	-0.0009
	(0.041)	(0.0002)	(0.0011)	(0.0022)	(0.0023)	(0.0006)
	0.0635	0.0759	0.5282	0.8577	0.8829	0.1283
School Share Absence Rate Above Median						
Yes	0.202	0.0012	0.0116	0.0008	-0.0023	-0.0002
	(0.028)	(0.0002)	(0.0007)	(0.0014)	(0.0014)	(0.0004)
	16,689	16,689	16,689	9,576	9,565	15,538
No	0.168	0.0009	0.0114	0.0002	-0.0038	0.0002
	(0.029)	(0.0002)	(0.0008)	(0.0016)	(0.0016)	(0.0004)
	12,664	12,664	12,664	7,264	7,275	11,774
Difference in Means	0.034	0.0003	0.0002	0.0007	0.0015	-0.0004
	(0.040)	(0.0002)	(0.0011)	(0.0021)	(0.0022)	(0.0006)
	0.3923	0.2646	0.8240	0.7604	0.4966	0.5287
School Share Changed School Above Median						
Yes	0.199	0.0011	0.0119	0.0001	-0.0030	-0.0002
	(0.025)	(0.0001)	(0.0006)	(0.0013)	(0.0013)	(0.0003)
	20,403	20,403	20,403	11,626	11,642	18,986
No	0.148	0.0009	0.0106	0.0027	-0.0010	0.0005
	(0.035)	(0.0002)	(0.0010)	(0.0019)	(0.0020)	(0.0005)
	8,950	8,950	8,950	5,214	5,198	8,326
Difference in Means	0.051	0.0002	0.0013	-0.0026	-0.0020	-0.0007
	(0.043)	(0.0003)	(0.0012)	(0.0023)	(0.0023)	(0.0006)
	0.2300	0.4335	0.2802	0.2646	0.4004	0.2142
School Mean Days Absent Z (in SD) Chg.						
Q1	0.089	0.0005	0.0000	0.0031	0.0004	0.0003
	(0.059)	(0.0004)	(0.0000)	(0.0036)	(0.0033)	(0.0009)
	3,423	3,423	3,423	1,799	1,811	3,085
Q2	0.073	0.0003	0.0000	0.0013	-0.0004	0.0014
	(0.054)	(0.0003)	(0.0000)	(0.0030)	(0.0030)	(0.0008)
	3,825	3,825	3,825	1,965	1,976	3,511
Q3	0.052	-0.0000	0.0000	0.0009	-0.0042	0.0005
	(0.061)	(0.0004)	(0.0000)	(0.0035)	(0.0036)	(0.0010)
	3,021	3,021	3,021	1,532	1,526	2,726
Q4	0.238	0.0013	0.0000	0.0011	-0.0032	0.0000
	(0.069)	(0.0004)	(0.0000)	(0.0033)	(0.0034)	(0.0009)
	3,422	3,422	3,422	1,854	1,871	3,099
Difference in Means	0.149	0.0009	0.0000	-0.0020	-0.0037	-0.0003
	(0.091)	(0.0005)	(0.0000)	(0.0049)	(0.0047)	(0.0012)
	0.1000	0.1106	.	0.6866	0.4373	0.8264

This table conducts a heterogeneity analysis by repeating the main analysis from Table 1 for subgroups. Unit of observation is a student school year. Outcomes are listed in columns. Rows index the characteristics and levels defining the subsamples among which the heterogeneity analysis is conducted. Each cell in a characteristic-level row reports the coefficient on continuous commute distance (in miles) from a separate regression of the column-enumerated outcome on commute distance and the main covariate specification from Table 1 for the subsample defined by the characteristic-level row. Standard errors clustered by family group are given in parentheses. Sample sizes are given in the third row of each cell. Difference in Means row for each characteristic gives the difference in coefficients, standard errors of the differences (in parentheses), and p-values (in the third row). For binary characteristics, the comparison is between the two levels; for ordered categorical variables, the comparison is between the highest and lowest levels.

Table A.15: School Quality Results

		Commute Distance				Out-of-Borough			
	Mean	Main	S+S FE	Stud. FE	Full	Main	S+S FE	Stud. FE	Full
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
School Mean Days Absent Chg.	0.5931	-0.0197	-0.0147	-0.0817	-0.1481	-0.0970	-0.2012	-0.9049	-1.0682
	(5.1921)	(0.0077)	(0.0070)	(0.0309)	(0.0456)	(0.1063)	(0.0999)	(0.4473)	(0.6795)
	13,691	13,691	13,556	1,743	1,297	13,691	13,556	1,743	1,297
		[-33.5]	[-24.9]	[-138.7]	[-251.4]	[-16.4]	[-33.9]	[-152.6]	[-180.1]
School Absence Rate Chg.	0.0039	-0.0002	-0.0001	-0.0006	-0.0012	-0.0007	-0.0015	-0.0069	-0.0087
	(0.0350)	(0.0001)	(0.0000)	(0.0002)	(0.0003)	(0.0007)	(0.0007)	(0.0029)	(0.0046)
	13,691	13,691	13,556	1,743	1,297	13,691	13,556	1,743	1,297
		[-40.0]	[-28.0]	[-158.9]	[-309.6]	[-17.4]	[-38.2]	[-180.2]	[-226.4]
School Promotion Rate Chg.	-0.0027	0.0002	0.0002	0.0007	0.0021	-0.0002	0.0010	0.0046	0.0113
	(0.0471)	(0.0001)	(0.0001)	(0.0003)	(0.0005)	(0.0010)	(0.0010)	(0.0032)	(0.0067)
	13,691	13,691	13,556	1,743	1,297	13,691	13,556	1,743	1,297
		[-81.2]	[-62.0]	[-250.5]	[-774.1]	[7.2]	[-38.1]	[-167.5]	[-412.6]
School Share Homeless Chg.	0.0082	-0.0001	0.0000	-0.0008	-0.0028	0.0015	0.0015	-0.0064	-0.0243
	(0.0587)	(0.0001)	(0.0001)	(0.0004)	(0.0006)	(0.0011)	(0.0011)	(0.0051)	(0.0074)
	13,691	13,691	13,556	1,743	1,297	13,691	13,556	1,743	1,297
		[-8.8]	[3.6]	[-101.6]	[-341.5]	[18.6]	[18.8]	[-78.1]	[-297.7]
School Share Poor Chg.	0.0063	-0.0003	-0.0001	-0.0002	-0.0004	-0.0022	0.0011	-0.0008	-0.0106
	(0.1367)	(0.0002)	(0.0002)	(0.0006)	(0.0016)	(0.0024)	(0.0024)	(0.0087)	(0.0215)
	13,691	13,691	13,556	1,743	1,297	13,691	13,556	1,743	1,297
		[-45.4]	[-11.9]	[-31.3]	[-58.7]	[-34.7]	[17.4]	[-12.7]	[-168.6]
School Enrollment Chg.	-22.9673	1.6574	0.8296	4.0278	3.7989	-7.4782	0.1103	66.3128	106.3248
	(422.9598)	(0.5550)	(0.5480)	(2.2226)	(3.7186)	(7.3959)	(7.4916)	(29.6123)	(51.6140)
	13,691	13,691	13,556	1,743	1,297	13,691	13,556	1,743	1,297
		[-72.7]	[-36.4]	[-176.6]	[-166.5]	[32.6]	[-0.5]	[-288.7]	[-462.9]
School Proficient Rate Chg.	-0.0105	0.0005	0.0002	0.0012	0.0036	0.0013	0.0026	0.0040	0.0464
	(0.1202)	(0.0002)	(0.0002)	(0.0007)	(0.0013)	(0.0025)	(0.0022)	(0.0097)	(0.0184)
	13,604	13,604	13,473	1,726	1,280	13,604	13,473	1,726	1,280
		[-46.2]	[-15.1]	[-114.3]	[-340.8]	[-12.2]	[-24.3]	[-38.1]	[-440.8]

Outcomes are listed in rows. Analytical specifications are indexed by column. Unit of observation is a student school year. Column 1 gives outcome means (standard deviations in parentheses). Each cell in Columns 2–5 reports the coefficient on continuous school-shelter commute distance (in miles) from a separate OLS linear regression of the row-indexed outcome on commute distance (i.e., treatment) and the covariates described at the bottom of the table. Each cell in Columns 6–9 reports analogous results for treatment defined as an indicator for out-of-school-borough shelter placement. Standard errors clustered by family group are given in parentheses. Sample sizes are given below standard errors. Percent changes from outcome means are given in brackets. All results are for the main sample, though only a subset of observations contribute to identification in the specifications with school, shelter, and, especially, student fixed effects.

Table A.16: Mediating Assocation of School Changes with Commute Distance

	Days Absent (1)	Absence Rate (2)	ELA Standardized (3)	Math Standardized (4)	Promoted (5)
Commute Distance (miles)	0.212	0.0012	0.0018	-0.0005	-0.0000
	(0.026)	(0.0001)	(0.0014)	(0.0014)	(0.0003)
School Change	2.677	0.0237	-0.0519	-0.0810	-0.0283
	(0.404)	(0.0023)	(0.0228)	(0.0221)	(0.0059)
School Change × Commute Distance (miles)	-0.093	-0.0006	-0.0007	-0.0021	0.0006
	(0.038)	(0.0002)	(0.0020)	(0.0020)	(0.0005)
N	29,353	29,353	16,840	16,840	27,312
Student and Family Covariates	Yes	Yes	Yes	Yes	Yes
Prior School Year Covs.	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes

This table explores school changes as a potential mediator for the effects of commute distance on student outcomes. Outcomes are listed in columns. Each column is a separate regression of the given outcome on continuous commute distance (in miles) interacted with an indicator for change of school, controlling for main covariates. Rows report the coefficients, respectively, of commute distance, school change, and their interaction. Results should be interpreted as associations, since school changes following shelter entry are potentially endogenous. Standard errors clustered by family group are given in parentheses. Sample sizes are given below standard errors.

Table A.17: Homelessness Outcomes

		Commute Distance				Out-of-Borough			
	Mean	Main	S+S FE	Stud. FE	Full	Main	S+S FE	Stud. FE	Full
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Length of Stay during School Year	158.227	-0.323	-0.300	-0.323	-0.496	-3.679	-3.895	-6.383	-7.987
	(102.264)	(0.092)	(0.096)	(0.318)	(0.390)	(1.173)	(1.206)	(4.118)	(4.789)
	29,353	29,353	29,263	5,174	4,852	29,353	29,263	5,174	4,852
		[-2.0]	[-1.8]	[-2.0]	[-3.0]	[-2.3]	[-2.5]	[-4.0]	[-5.0]
Log Length of Stay during School Year	4.7217	-0.0043	-0.0040	-0.0056	-0.0089	-0.0512	-0.0510	-0.1005	-0.1494
	(1.0338)	(0.0009)	(0.0009)	(0.0033)	(0.0040)	(0.0115)	(0.0120)	(0.0428)	(0.0504)
	29,353	29,353	29,263	5,174	4,852	29,353	29,263	5,174	4,852
		[-0.9]	[-0.8]	[-1.1]	[-1.8]	[-1.1]	[-1.1]	[-2.1]	[-3.2]
Length of Stay	457.451	-1.166	-1.136	-3.032	-2.521	-21.098	-21.805	-43.125	-34.211
	(407.282)	(0.532)	(0.549)	(1.127)	(1.403)	(6.731)	(6.933)	(14.652)	(18.202)
	29,353	29,353	29,263	5,174	4,852	29,353	29,263	5,174	4,852
		[-2.4]	[-2.4]	[-6.4]	[-5.3]	[-4.6]	[-4.8]	[-9.4]	[-7.5]
Log Length of Stay	5.6386	-0.0071	-0.0064	-0.0128	-0.0153	-0.0968	-0.0930	-0.2078	-0.2444
	(1.1626)	(0.0015)	(0.0016)	(0.0047)	(0.0056)	(0.0196)	(0.0202)	(0.0606)	(0.0693)
	29,353	29,353	29,263	5,174	4,852	29,353	29,263	5,174	4,852
		[-1.2]	[-1.1]	[-2.2]	[-2.6]	[-1.7]	[-1.6]	[-3.7]	[-4.3]
Homeless First School Year Post-Entry	0.7648	-0.0007	-0.0005	-0.0022	-0.0023	-0.0043	-0.0036	-0.0318	-0.0237
	(0.4241)	(0.0006)	(0.0006)	(0.0016)	(0.0017)	(0.0071)	(0.0074)	(0.0207)	(0.0223)
	27,378	27,378	27,290	4,652	4,322	27,378	27,290	4,652	4,322
		[-0.9]	[-0.7]	[-2.8]	[-2.9]	[-0.6]	[-0.5]	[-4.2]	[-3.1]
Homeless Second School Year Post-Entry	0.4885	-0.0001	0.0003	-0.0002	0.0029	-0.0003	-0.0003	0.0027	0.0211
	(0.4999)	(0.0008)	(0.0008)	(0.0021)	(0.0028)	(0.0098)	(0.0102)	(0.0274)	(0.0360)
	21,362	21,362	21,285	3,274	2,911	21,362	21,285	3,274	2,911
		[-0.2]	[0.6]	[-0.4]	[5.8]	[-0.1]	[-0.1]	[0.5]	[4.3]
Sample	Main	Main	Main	Stud FE	Stud FE	Main	Main	Stud FE	Stud FE
Student and Family Covariates	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prior School Year Covs.	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School and Shelter FE	No	No	Yes	No	Yes	No	Yes	No	Yes
Student FE	No	No	No	Yes	Yes	No	No	Yes	Yes

This table repeats Table 1 for homelessness outcomes. Outcomes are listed in rows. Analytical specifications are indexed by column. Unit of observation is a student school year. Column 1 gives outcome means (standard deviations in parentheses). Each cell in Columns 2–5 reports the coefficient on continuous school-shelter commute distance (in miles) from a separate OLS linear regression of the row-indexed outcome on commute distance (i.e., treatment) and the covariates described at the bottom of the table. Each cell in Columns 6–9 reports analogous results for treatment defined as an indicator for out-of-school-borough shelter placement. Standard errors clustered by family group are given in parentheses. Sample sizes are given below standard errors. Percent changes from outcome means are given in brackets. All results are for the main sample, though only a subset of observations contribute to identification in the specifications with school, shelter, and, especially, student fixed effects.

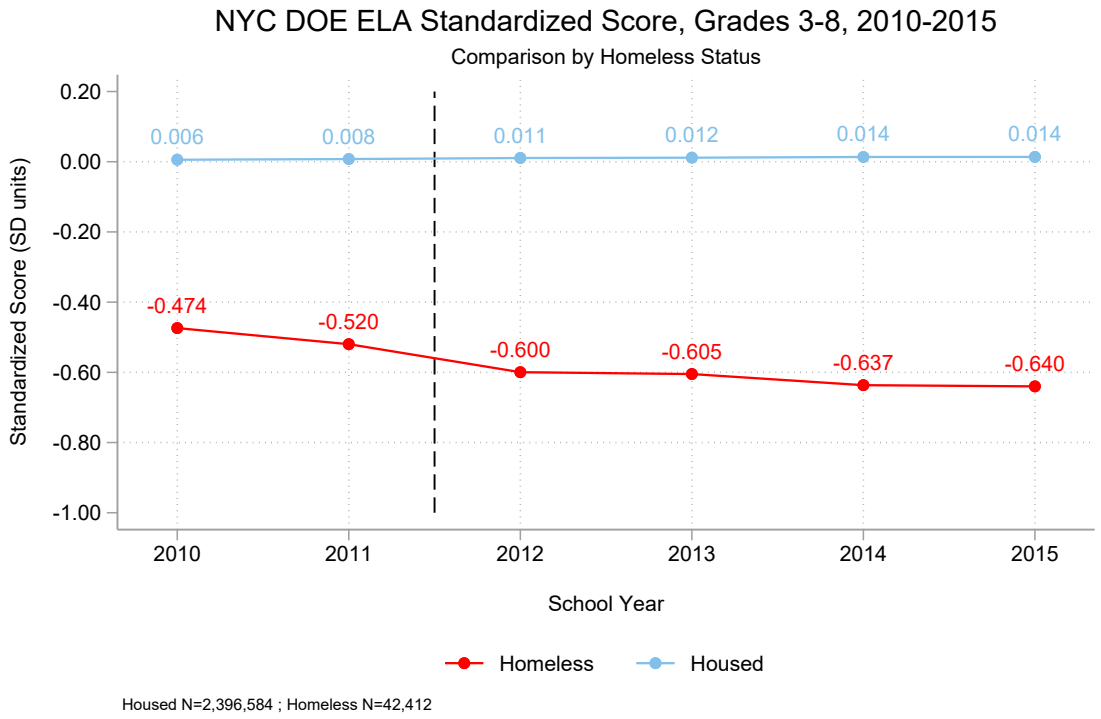
Table A.18: School-Shelter Commute Distance and Future Year Outcomes

	Year t+1			Year t+2		
	Mean	Main	S+S FE	Mean	Main	S+S FE
	(1)	(2)	(3)	(4)	(5)	(6)
Days Absent	26.075	0.015	0.014	26.803	-0.053	-0.054
	(22.468)	(0.023)	(0.024)	(25.829)	(0.030)	(0.031)
	27,187	27,187	27,100	21,211	21,211	21,133
Absence Rate	0.1550	-0.0000	-0.0000	0.1633	-0.0003	-0.0003
	(0.1364)	(0.0001)	(0.0001)	(0.1601)	(0.0002)	(0.0002)
	27,013	27,013	26,926	20,919	20,919	20,839
Changed School	0.3703	0.0042	0.0039	0.3074	0.0014	0.0013
	(0.4829)	(0.0006)	(0.0006)	(0.4614)	(0.0007)	(0.0007)
	27,378	27,378	27,290	21,362	21,362	21,285
ELA Standardized Score (SD units)	-0.5558	0.0013	0.0021	-0.5662	0.0002	0.0000
	(0.9223)	(0.0012)	(0.0013)	(0.9194)	(0.0014)	(0.0016)
	13,805	13,805	13,677	10,978	10,978	10,822
Math Standardized Score (SD units)	-0.6436	-0.0018	-0.0012	-0.6418	-0.0007	-0.0011
	(0.8679)	(0.0012)	(0.0013)	(0.8666)	(0.0014)	(0.0015)
	13,702	13,702	13,574	10,890	10,890	10,737
Promoted	0.9214	-0.0004	-0.0005	0.9107	0.0003	0.0003
	(0.2692)	(0.0003)	(0.0004)	(0.2851)	(0.0004)	(0.0004)
	20,984	20,984	20,904	15,819	15,819	15,723
Student and Family Covariates	No	Yes	Yes	No	Yes	Yes
Prior School Year Covs.	No	Yes	Yes	No	Yes	Yes
School and Shelter FE	No	No	Yes	No	No	Yes

This table repats the analysis of Table 1, Columns 1–3 for for outcomes in the school year following the shelter entry school year(year $t + 1$; Columns 1–3) and two years following shelter entry (year $t + 2$; Columns 4–6), where t is the school year of shelter entry. Outcomes are listed in rows. Analytical specifications are indexed by column. Treatment is continuous school-shelter commute distance (in miles). Unit of observation is a student school year. Each cell reports the coefficient on commute distance from a separate OLS linear regression of the row-indexed outcome on commute distance and the covariates described at the bottom of the table. Columns 1 and 4 give sample means. Columns 2 and 5 feature the main covariate specification. Columns 3 and 6 include school and shelter fixed effects. Standard errors clustered by family group are given in parentheses. Sample sizes are given below standard errors.

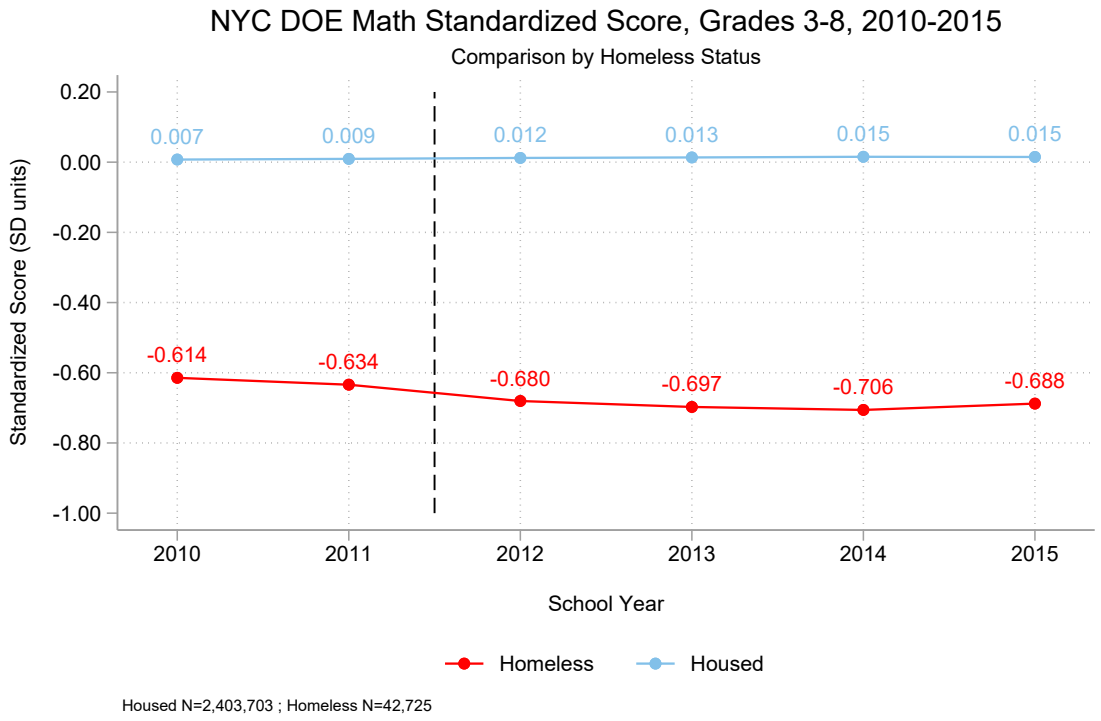
E Supplementary Figures

Figure A.1



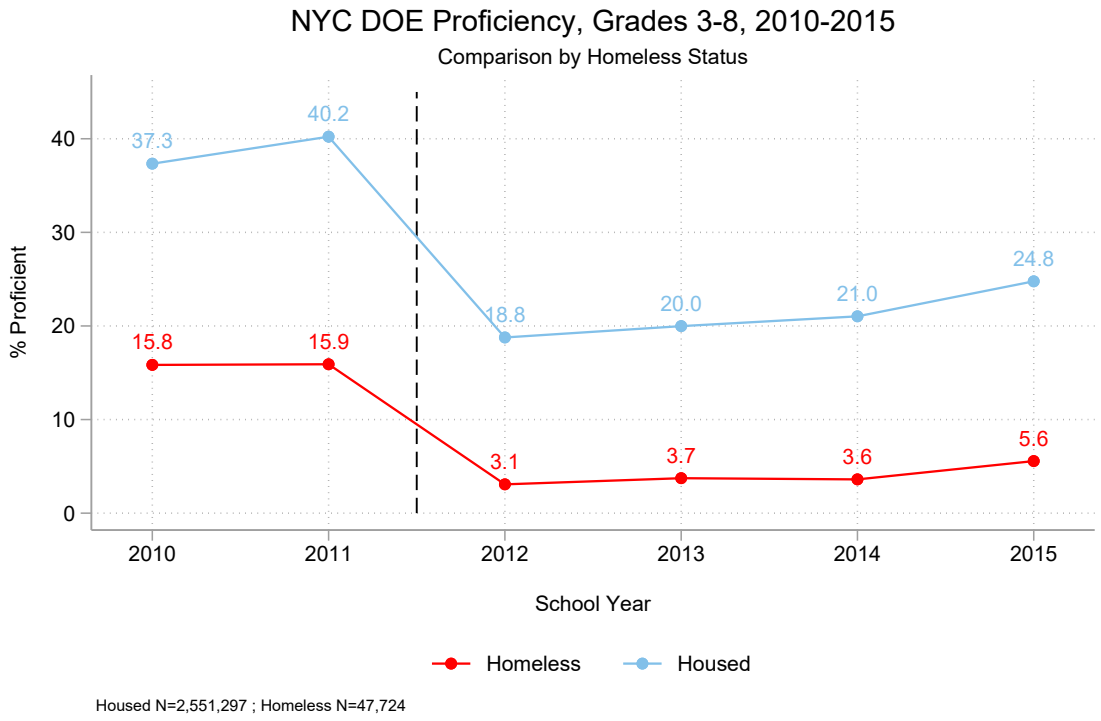
New York State grade 3–8 English Language Arts scaled test scores are normalized by year and grade within the NYC Department of Education student population; thus, point estimates are in standard deviation units. Homeless is defined as residing in Department of Homeless Services shelter; housed is defined as all other students. Sample covers the 2010–2015 school years among 3–8 public school students not in charter schools.

Figure A.2



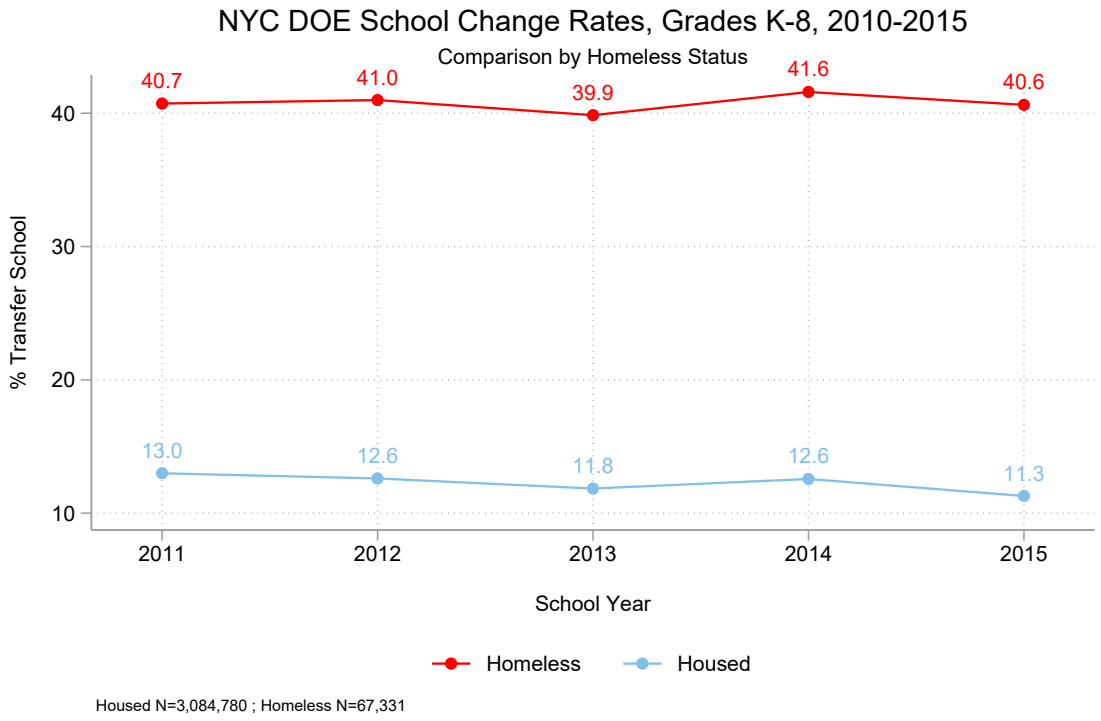
New York State grade 3–8 math scaled test scores are normalized by year and grade within the NYC Department of Education student population; thus, point estimates are in standard deviation units.Homeless is defined as residing in Department of Homeless Services shelter; housed is defined as all other students. Sample covers the 2010–2015 school years among 3–8 public school students not in charter schools.

Figure A.3



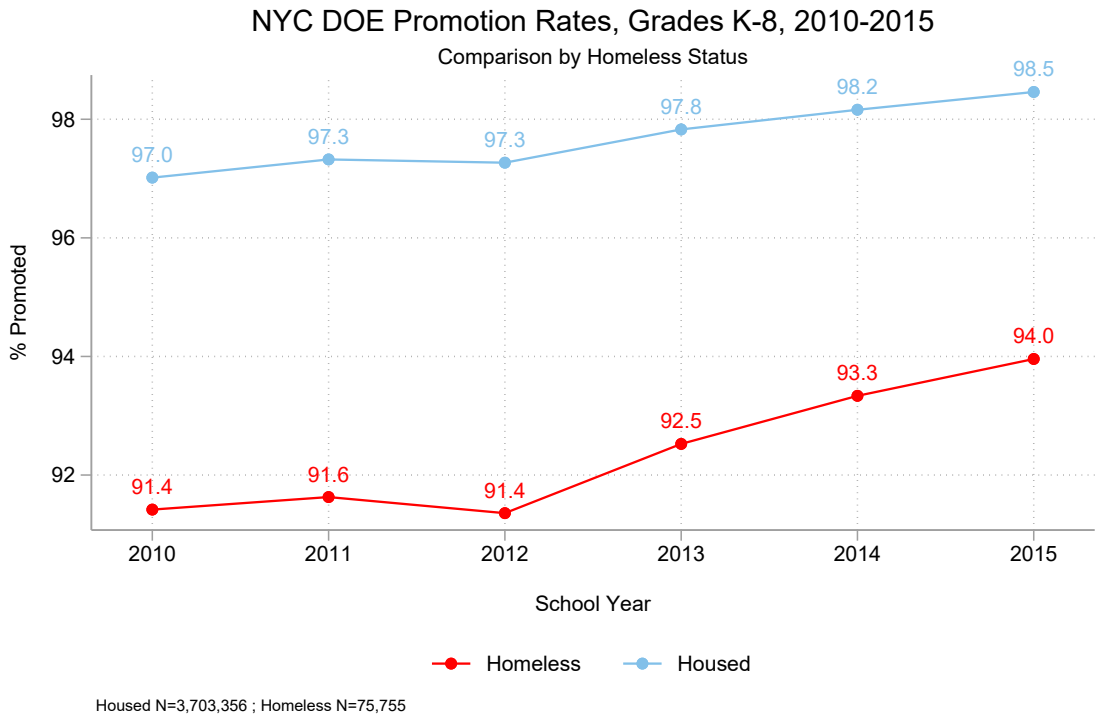
Proficient is defined as scoring level 3 or 4 on both English and math New York State tests. Students missing tests are considered not proficient. Homeless defined as residing in Department of Homeless Services shelter; housed is defined as all other students. Sample covers the 2010–2015 school years among 3–8 public school students not in charter schools. New Common Core testing standards were introduced in 2012.

Figure A.4



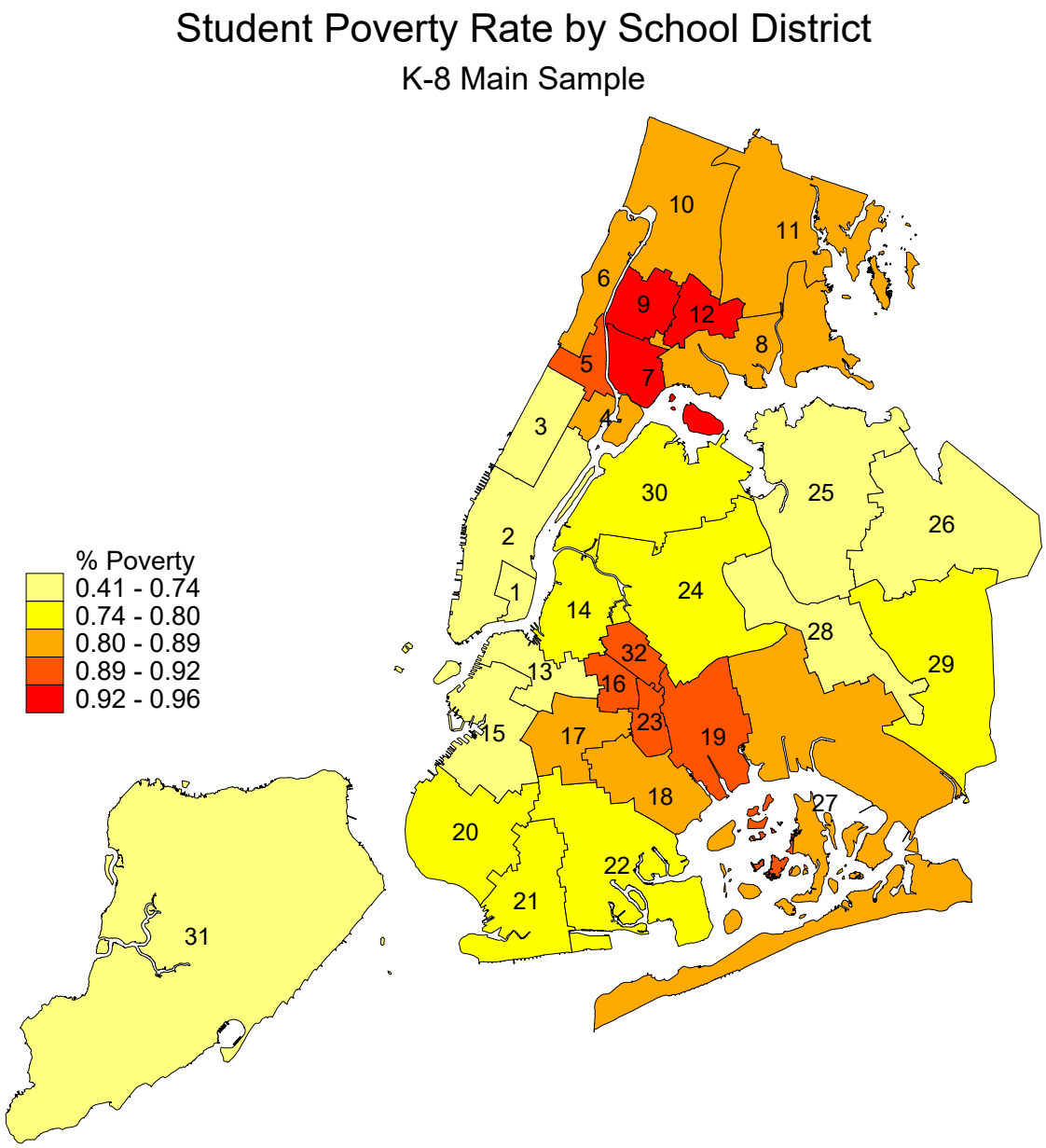
This figure gives the probabilities of students changing schools within a given school year. Homeless is defined as residing in Department of Homeless Services shelter; housed is defined as all other students. Sample covers the 2011–2015 school years among K–8 public school students not in charter schools. 2010 is excluded because 2009 homeless data is unavailable.

Figure A.5



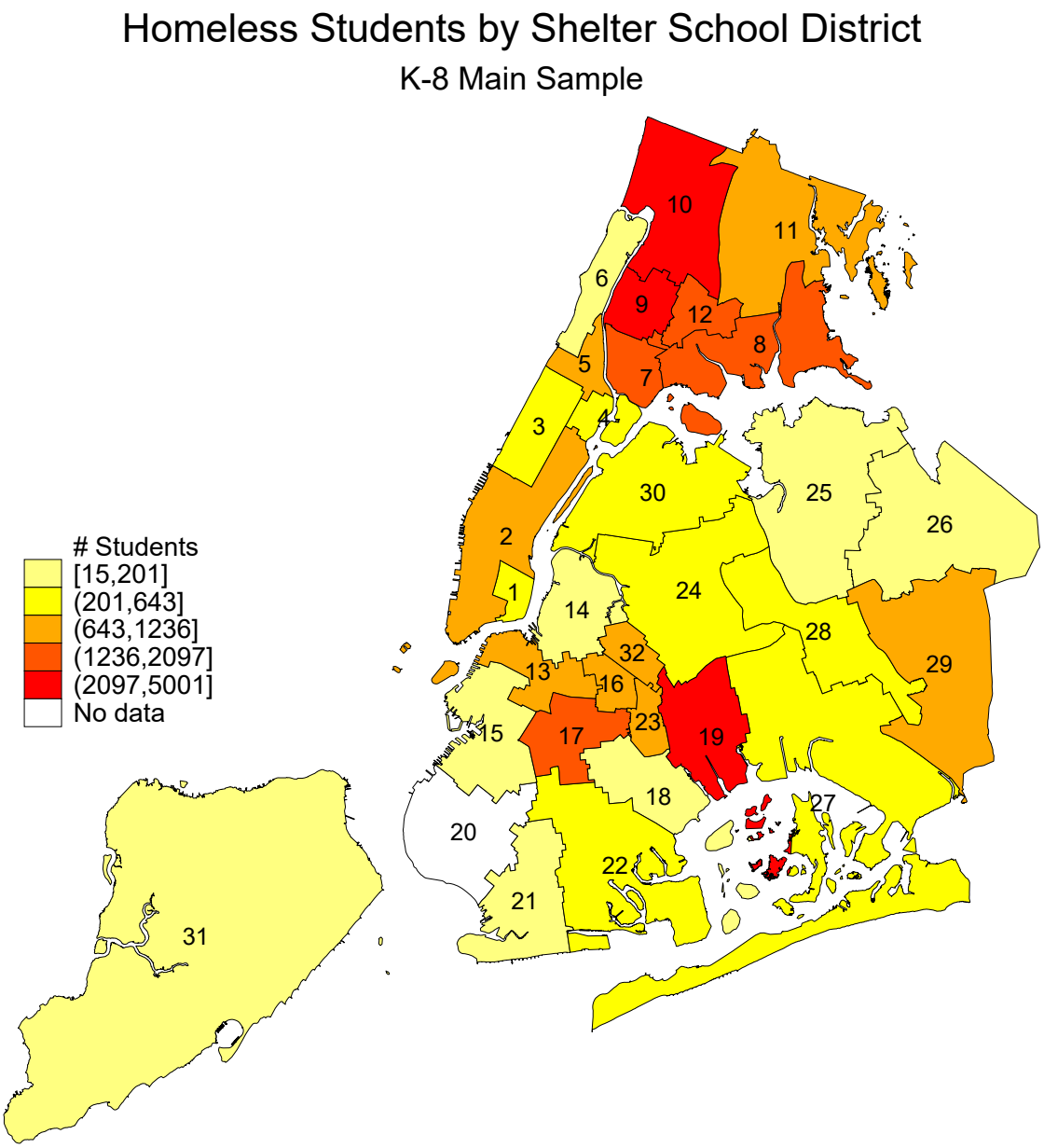
This figure gives the probabilities of students being promoted during each school year. Homeless is defined as residing in DHS shelter; housed is defined as all other students. Sample covers the 2010–2015 school years among 3–8 public school students not in charter schools.

Figure A.6



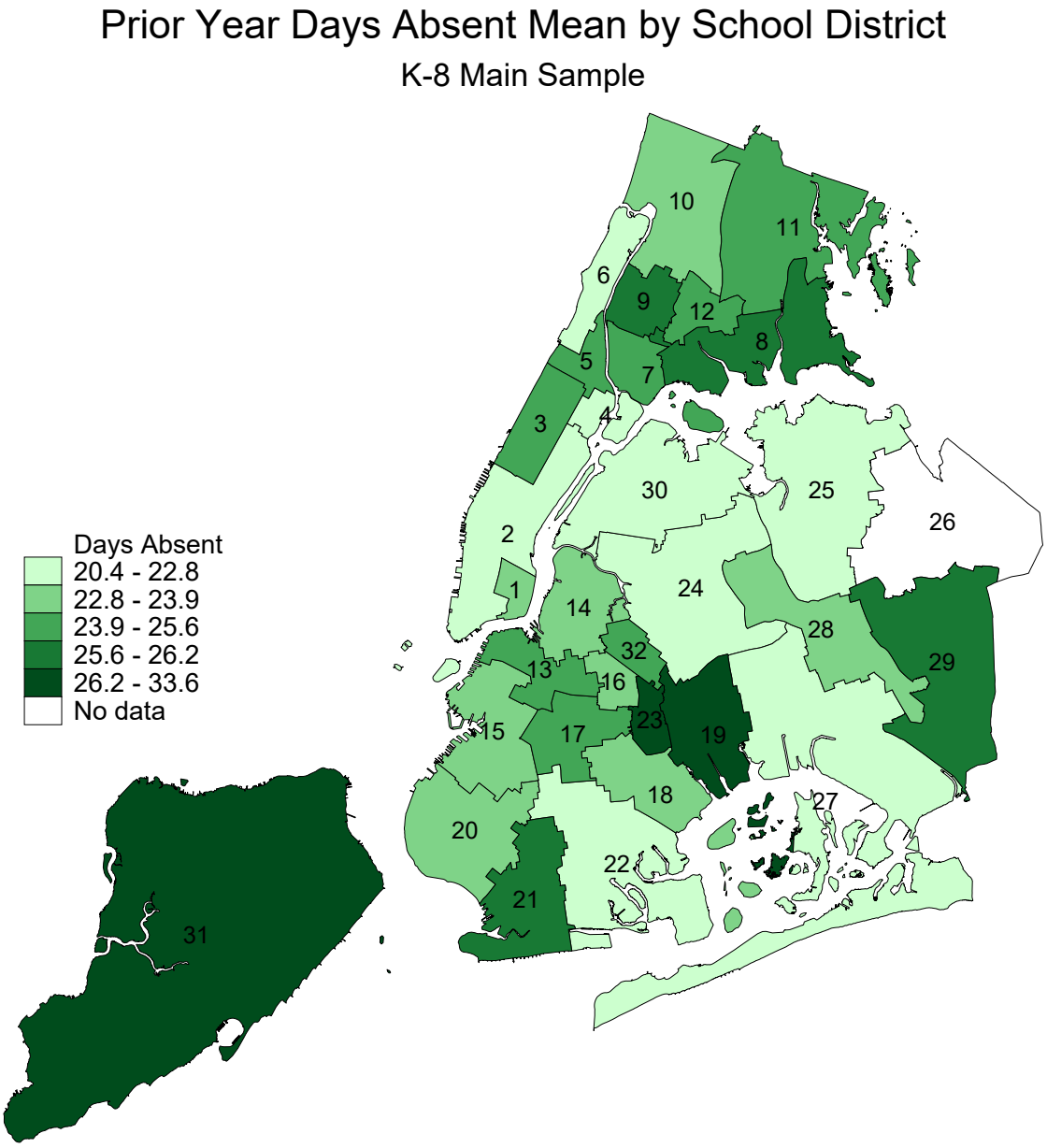
This figure depicts a choropleth map of mean poverty rates for the complete sample of primary schoolers by school district of origin, pooling school years 2010–2015. Poverty is defined as eligible for free or reduced-price lunch or public assistance. Limits of shading bins are set at 0, 25, 50, 75, 90, 100 percentiles of the mean poverty rate.

Figure A.7



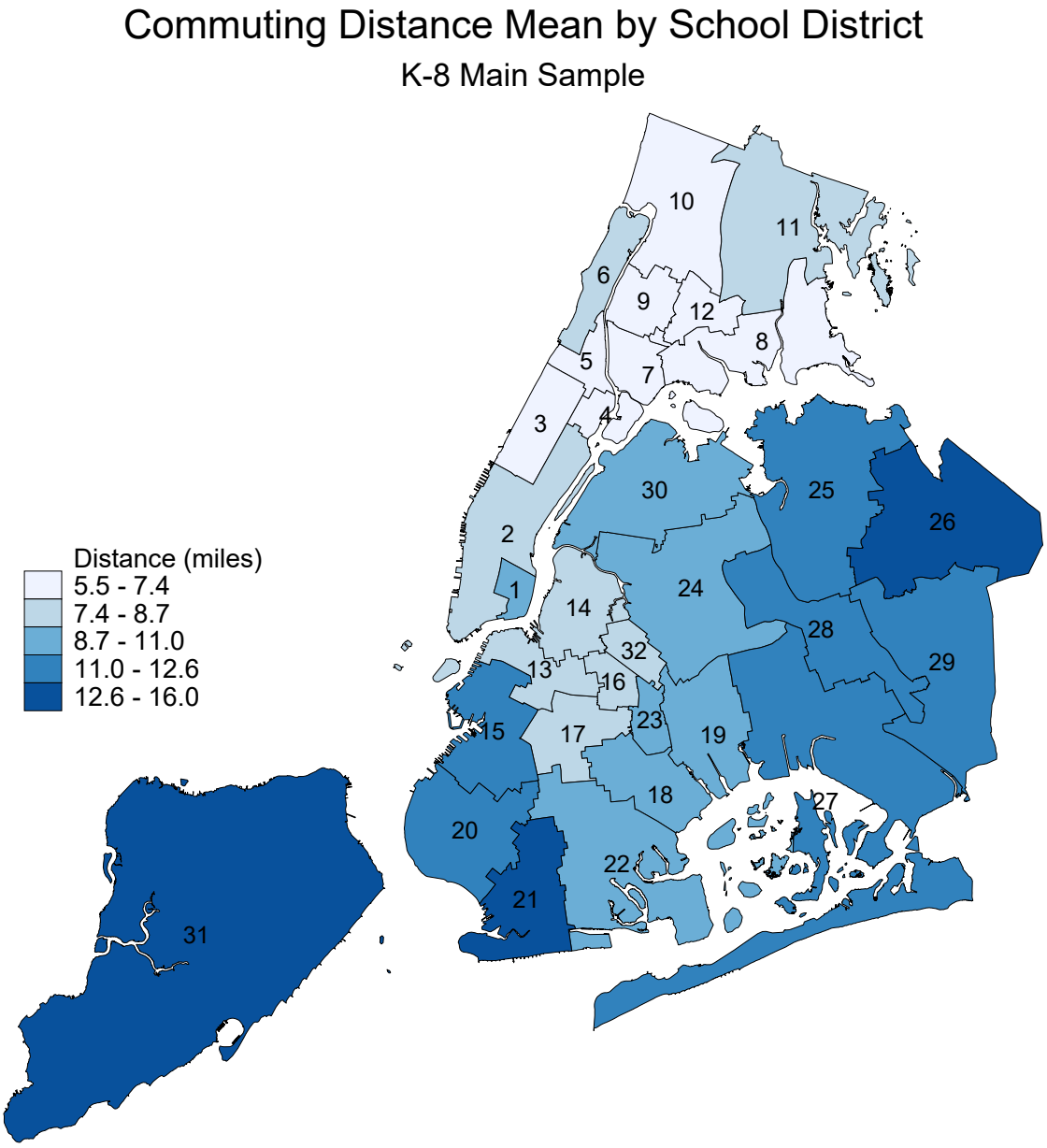
The figure depicts a heatmap of main sample homeless primary schoolers by school district of shelter, pooling school years 2010–2015. Limits of choropleth shading bins are set at 0, 25, 50, 75, 90, 100 percentiles of the homeless student counts.

Figure A.8



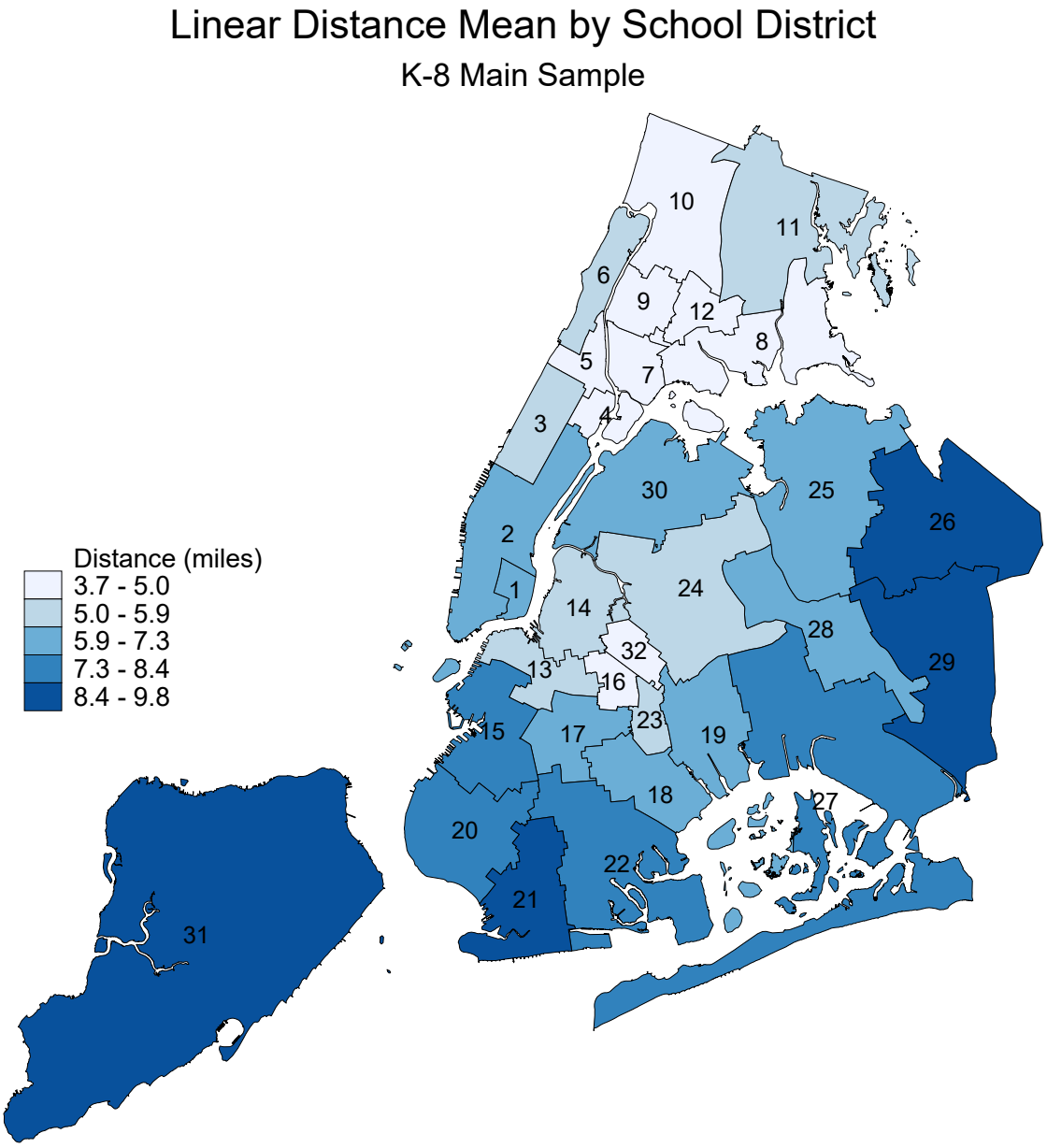
The figure depicts a choropleth map of mean days absent for the main sample of homeless primary schoolers in the school year prior to shelter entry by school district of origin, pooling school years 2010–2015. Limits of shading bins are set at 0, 25, 50, 75, 90, 100 percentiles of prior year days absent.

Figure A.9



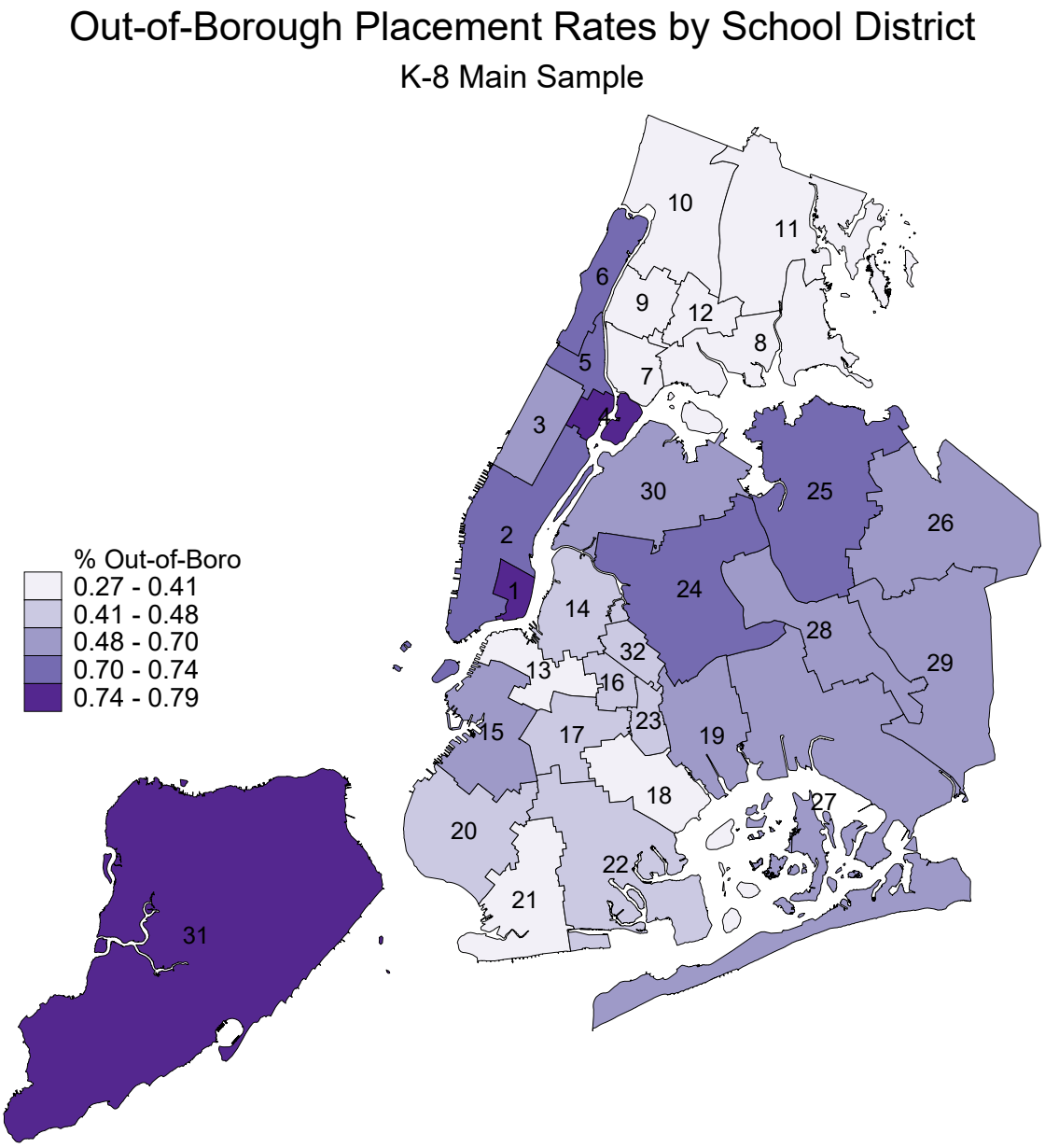
The figure depicts a choropleth map of mean commute distances for the main sample of homeless primary schoolers by school district of origin, pooling school years 2010–2015. Limits of shading bins are set at 0, 25, 50, 75, 90, 100 percentiles of commute distances.

Figure A.10



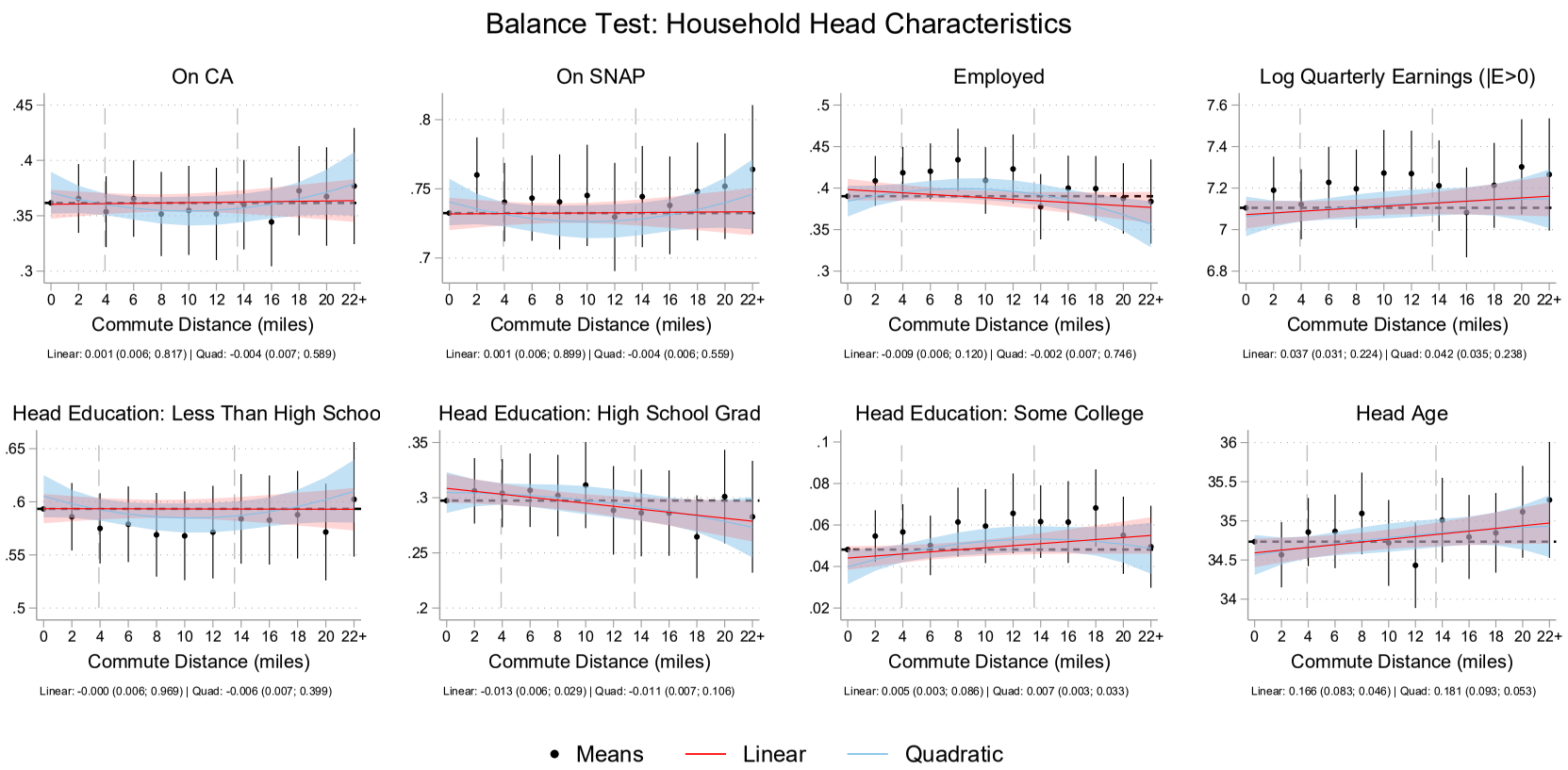
The figure depicts a choropleth map of mean linear school-shelter distances for the main sample of homeless primary schoolers by school district of origin, pooling school years 2010–2015. Limits of shading bins are set at 0, 25, 50, 75, 90, 100 percentiles of linear distance.

Figure A.11



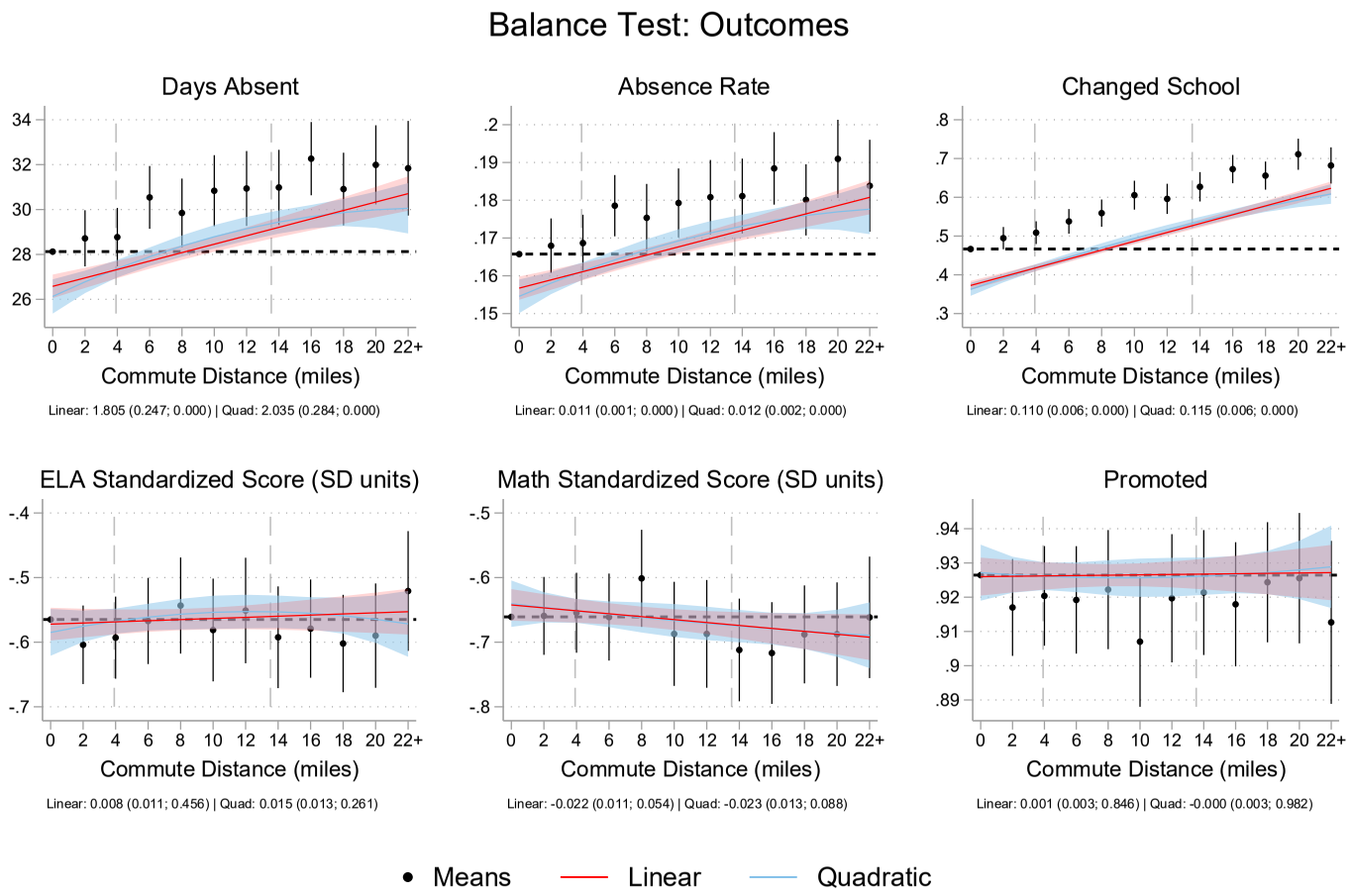
The figure depicts a choropleth map of mean out-of-borough shelter placement rate for the main sample of homeless primary schoolers by school district of origin, pooling school years 2010–2015. Limits of shading bins are set at 0, 25, 50, 75, 90, 100 percentiles of mean out-of-borough placement rate.

Figure A.12



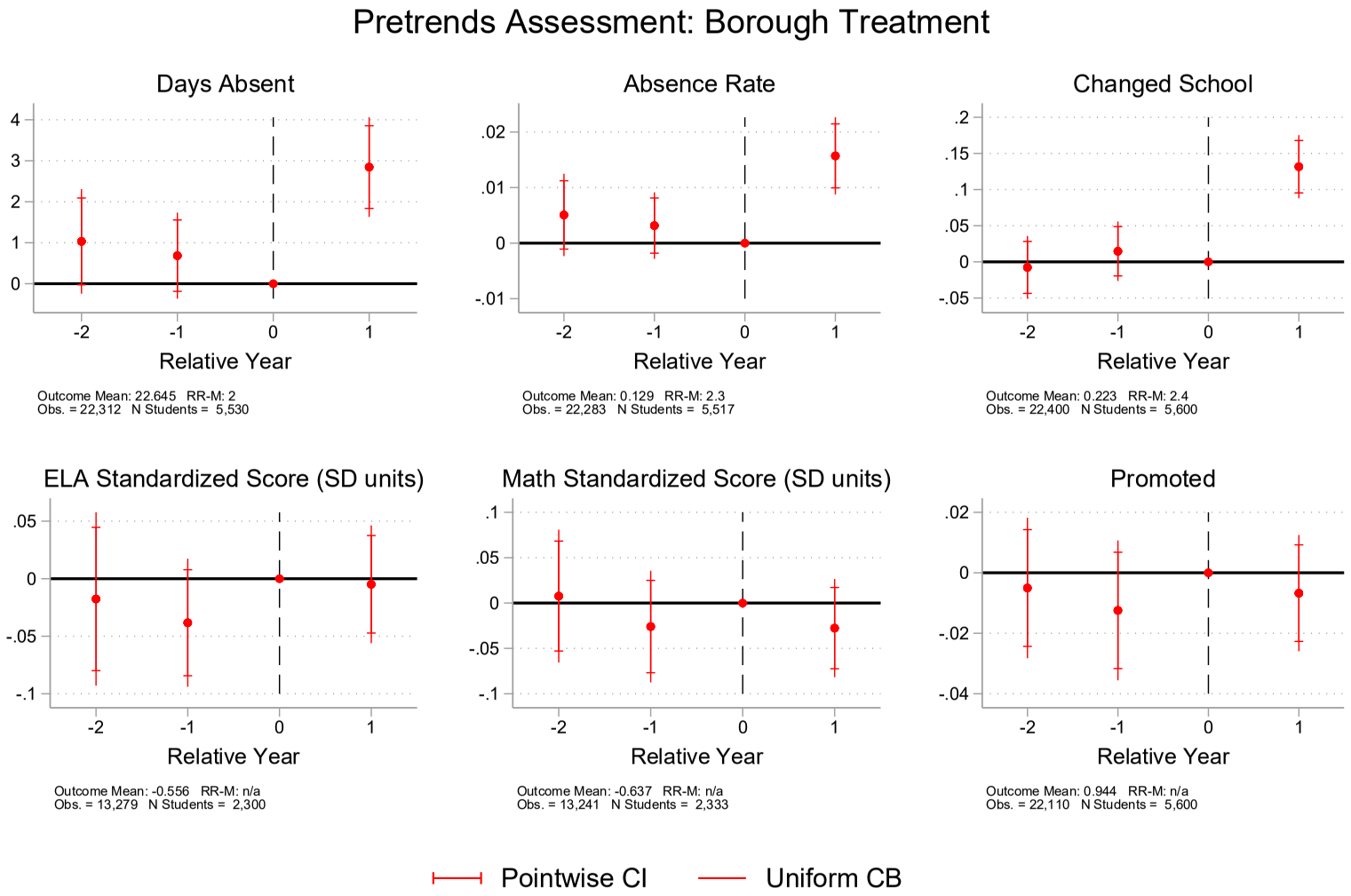
This figure presents a comprehensive balance test of household head characteristics for the main sample, split into one panel for each characteristic. Black dots and black vertical lines give covariate-adjusted means and 95 percent confidence intervals for the sample divided into 12 commute distance groups according to rounded-to-the-nearest-two-mile commute distances. (All students with commute distances 22 miles or greater are grouped together.) Estimates are obtained from separate linear regressions of each characteristic of interest on commute distance group fixed effects (omitting the 0-mile group as the baseline) and balance test covariates. Red and blue lines, respectively, give linear and quadratic fits (and shaded 95 percent confidence intervals) obtained from separate OLS regressions of each characteristic on continuous commute distance, controlling for balance test covariates. Dashed vertical gray lines give the in-borough (leftmost) and out-of-borough (rightmost) commute distance means. All standard errors are clustered by family group. Mean differences in characteristic means at the in- and out-of-borough commute distance means are reported under each panel, with standard errors and p-values presented in parentheses in that order.

Figure A.13



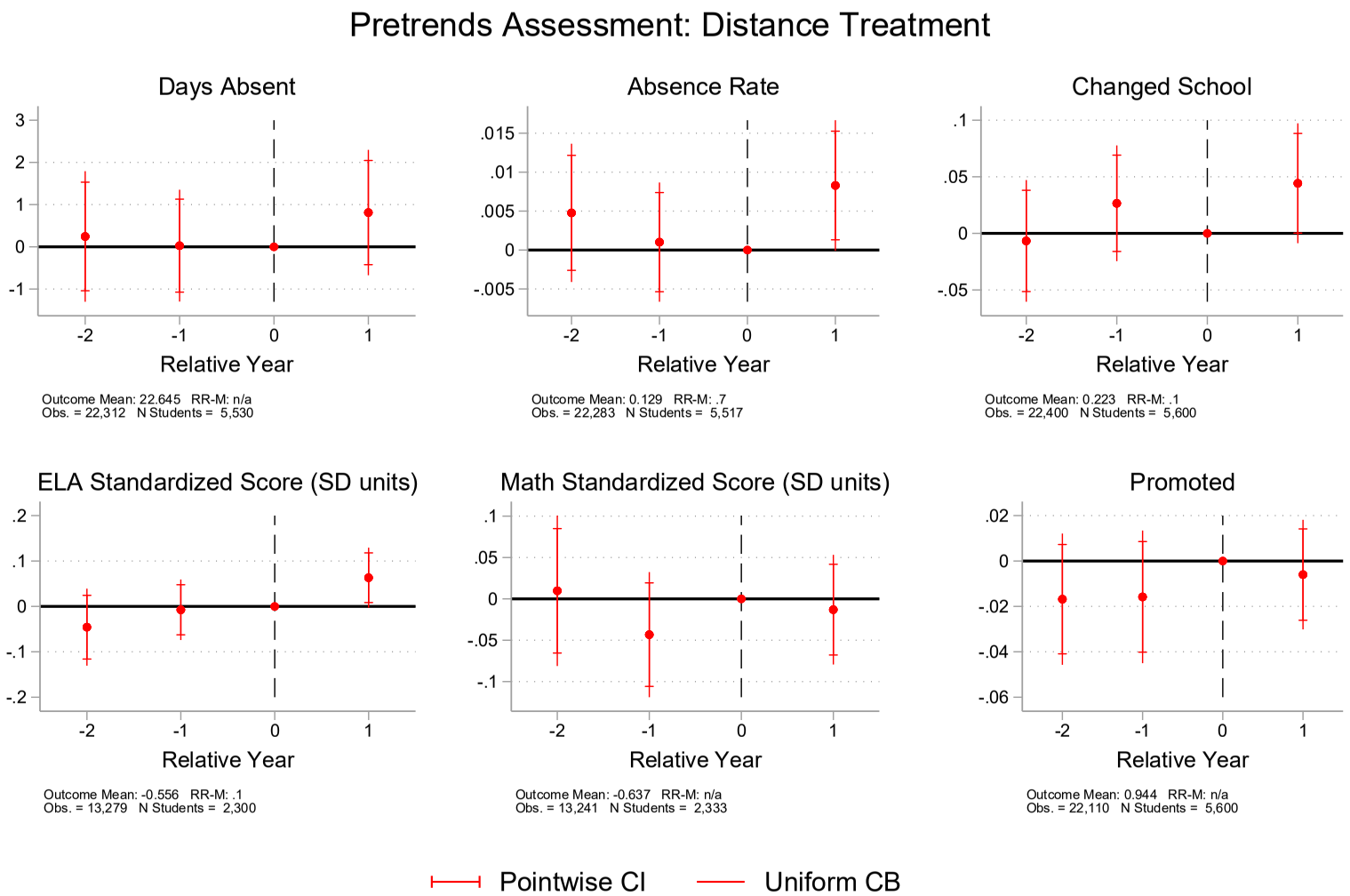
The figure presents a comprehensive balance test of student outcomes for the main sample, split into one panel for each characteristic. Black dots and black vertical lines give covariate-adjusted means and 95 percent confidence intervals for the sample divided into 12 commute distance groups according to rounded-to-the-nearest-two-mile commute distances. (All students with commute distances 22 miles or greater are grouped together.) Estimates are obtained from separate linear regressions of each characteristic of interest on commute distance group fixed effects (omitting the 0-mile group as the baseline) and balance test covariates. Red and blue lines, respectively, give linear and quadratic fits (and shaded 95 percent confidence intervals) obtained from separate OLS regressions of each characteristic on continuous commute distance, controlling for balance test covariates. Dashed vertical gray lines give the in-borough (leftmost) and out-of-borough (rightmost) commute distance means. All standard errors are clustered by family group. Mean differences in characteristic means at the in- and out-of-borough commute distance means are reported under each panel, with standard errors and p-values presented in parentheses in that order.

Figure A.14



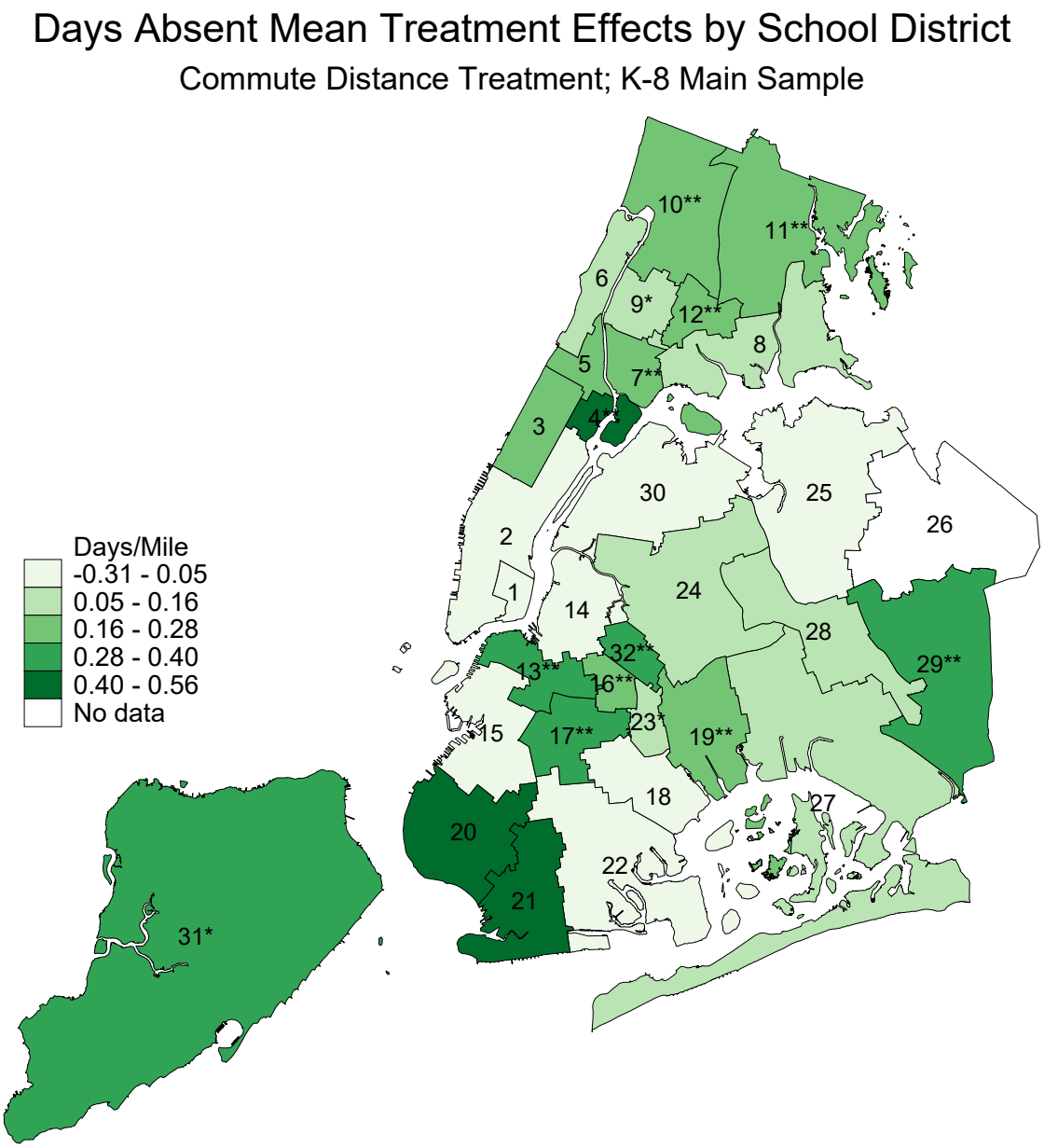
This figure assesses the credibility of the parallel trends assumption, with out-of-borough shelter assignment as the treatment. Point estimates and 95 percent confidence intervals are obtained from separate two-way fixed effects OLS linear regressions of each outcome (denoted by panel) on relative year interacted with an indicator for treatment group membership, controlling for year and student fixed effects and clustering standard errors by family group. Pointwise CI's are indicated by bars; uniform confidence bands are indicated by vertical lines. Outcome mean, Rambachan and Roth (2023) $\overline{M}$ statistic, sample size, and unique number of students reported below each panel.

Figure A.15



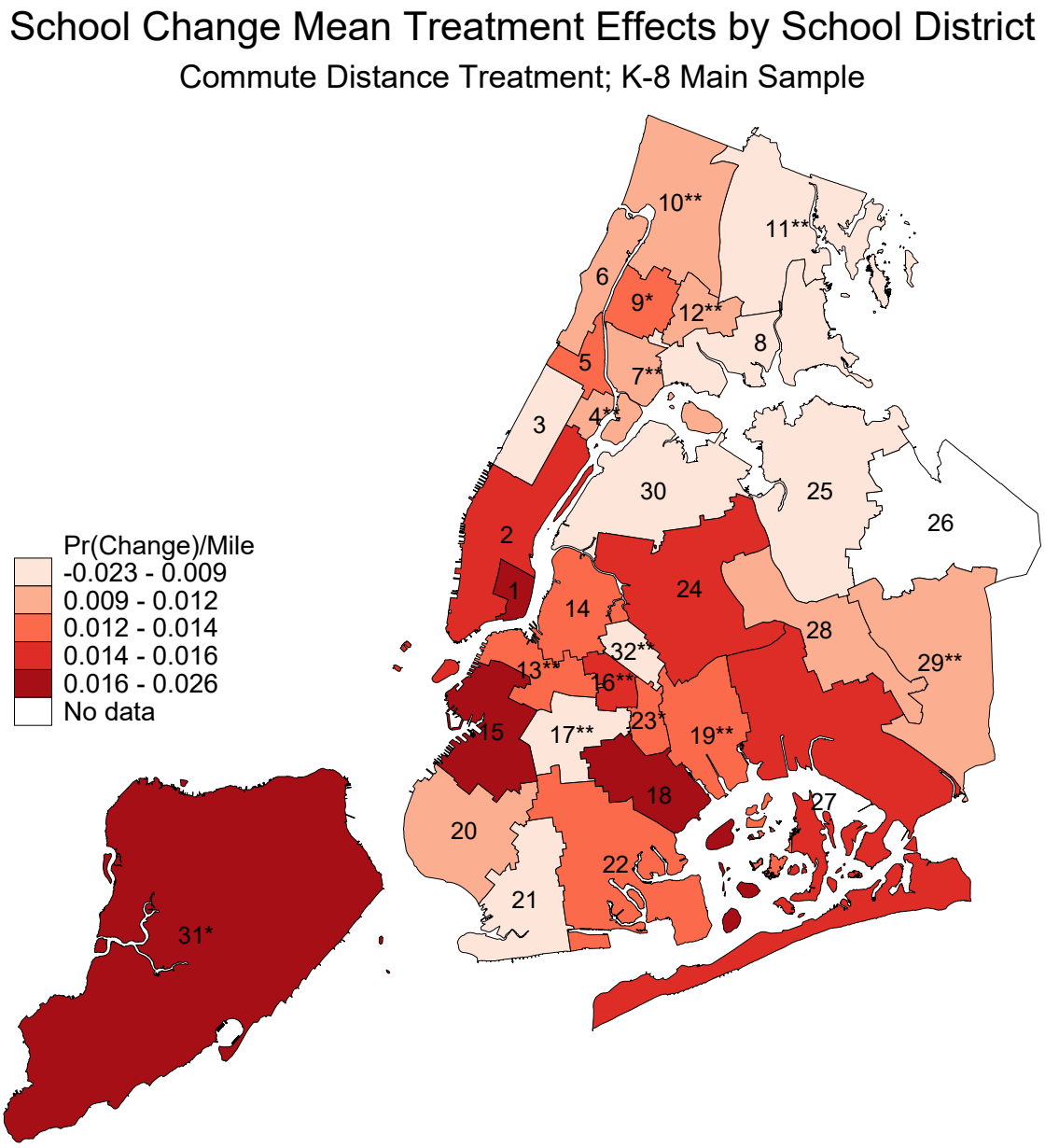
This figure assesses the credibility of the parallel trends assumption, with change linear school-shelter distance upon shelter entry as the treatment; for purposes of this figure, students whose absolute value of linear distance change is greater than one mile are considered the treatment group, and those with smaller changes compromise the control group. Point estimates and 95 percent confidence intervals are obtained from separate two-way fixed effects OLS linear regressions of each outcome (denoted by panel) on relative year interacted with an indicator for treatment group membership, controlling for year and student fixed effects and clustering standard errors by family group. Pointwise CI's are indicated by bars; uniform confidence bands are indicated by vertical lines. Outcome mean, Rambachan and Roth (2023) $\overline{M}$ statistic, sample size, and unique number of students reported below each panel.

Figure A.16



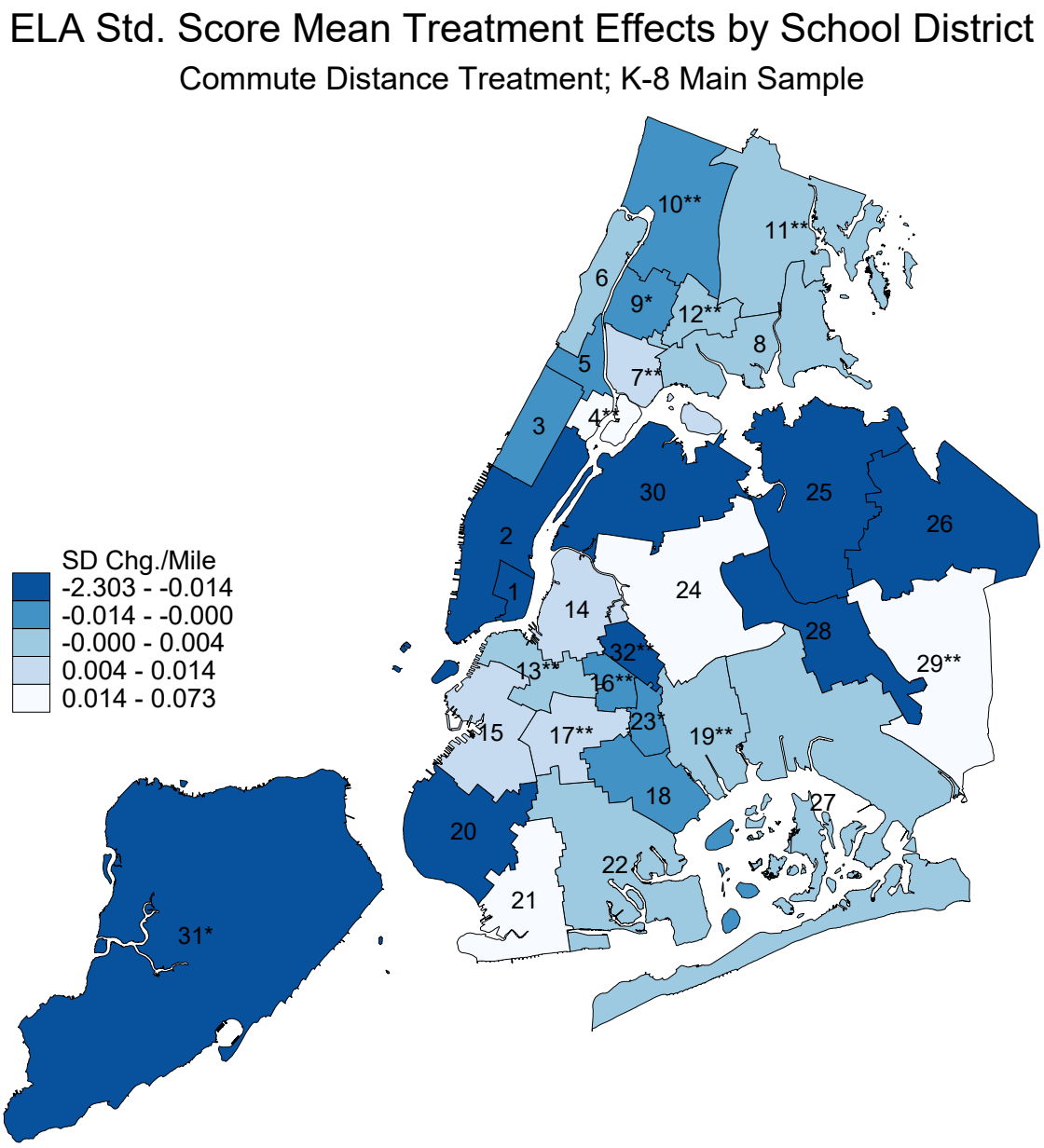
The figure depicts a choropleth map of average marginal effects of commute distance on days absent (in days per mile) for main sample homeless primary schoolers by school district of origin, pooling school years 2010–2015. Estimates are obtained from separate OLS linear regressions, within each school district, of days absent on commute distance, controlling for main covariates. Limits of shading bins are set at 0, 25, 50, 75, 90, 100 percentiles of effect (coefficient) distribution. $^{*} * p < 0.05,^{*} p < 0.10$.

Figure A.17



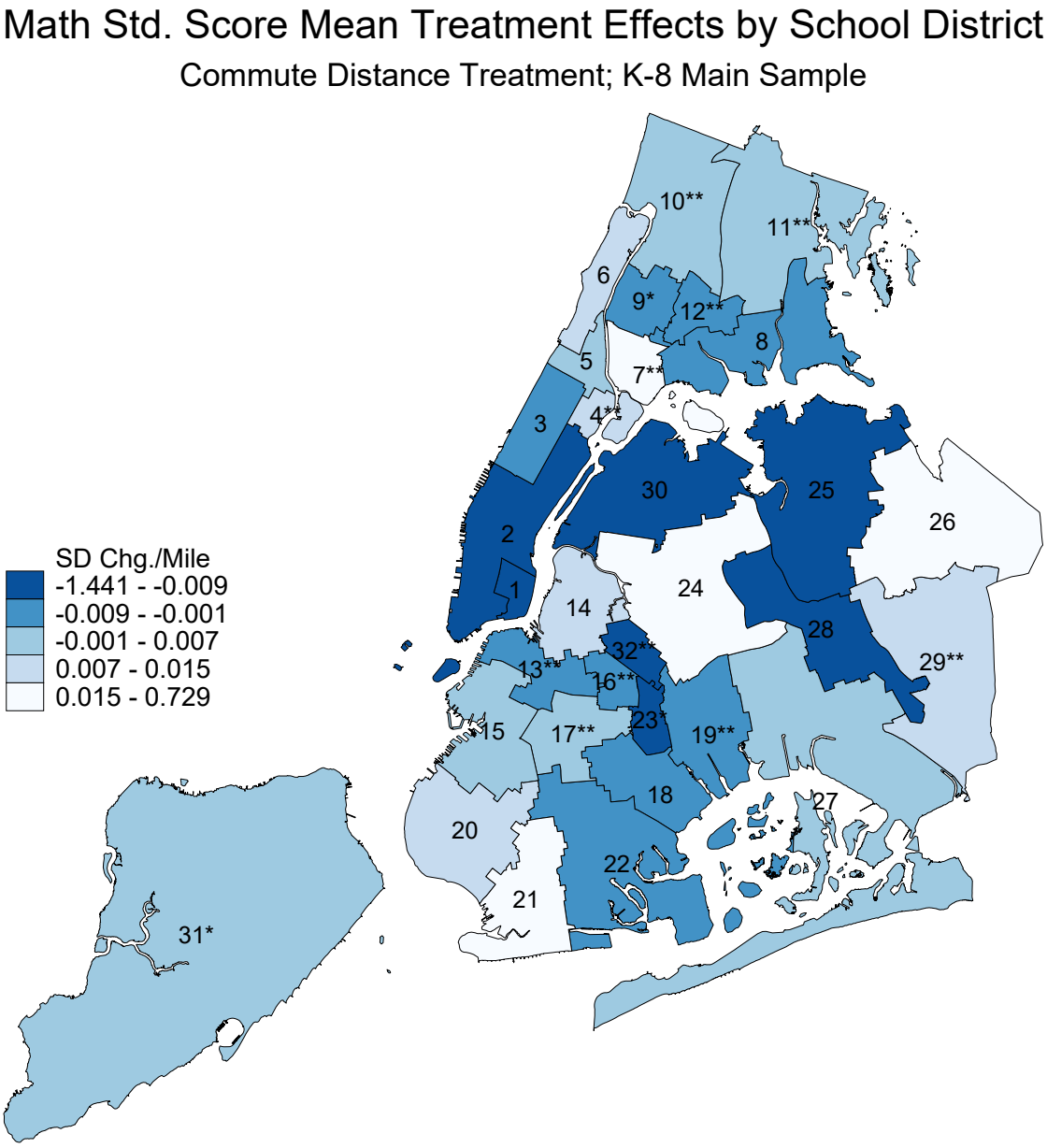
The figure depicts a choropleth map of average marginal effects of commute distance on the probability of school change for main sample homeless primary schoolers by school district of origin, pooling school years 2010–2015. Estimates are obtained from separate OLS linear regressions, within each school district, of days absent on commute distance, controlling for main covariates. Limits of shading bins are set at 0, 25, 50, 75, 90, 100 percentiles of effect (coefficient) distribution. $^{*} * p < 0.05,^{*} p < 0.10$.

Figure A.18



The figure depicts a choropleth map of average marginal effects of commute distance on normalized 3–8 grade NYS test scores (in standard deviations per mile) for main sample 3–8 grade homeless primary schoolers by school district of origin, pooling school years 2010–2015. Estimates are obtained from separate OLS linear regressions, within each school district, of normalized test scores on commute distance, controlling for main covariates. Limits of shading bins are set at 0, 25, 50, 75, 90, 100 percentiles of effect (coefficient) distribution. $^{*} * p < 0.05,^{*} p < 0.10$.

Figure A.19



The figure depicts a choropleth map of average marginal effects of commute distance on normalized 3–8 grade NYS test scores (in standard deviations per mile) for main sample 3–8 grade homeless primary schoolers by school district of origin, pooling school years 2010–2015. Estimates are obtained from separate OLS linear regressions, within each school district, of normalized test scores on commute distance, controlling for main covariates. Limits of shading bins are set at 0, 25, 50, 75, 90, 100 percentiles of effect (coefficient) distribution. $^{*} * p < 0.05,^{*} p < 0.10$.