Commuting to educational opportunity? School choice effects of mass transit expansion in Mexico City

Andrew Dustan*
Department of Economics
Vanderbilt University

Diana Ngo
Department of Economics
Occidental College

January 2018

Abstract

School choice policies aim to increase educational access by weakening the link between a student's residence and his choice set, but long commutes and other barriers may constrain families from selecting otherwise-desirable schools. Leveraging a mass transit expansion in Mexico City's suburbs as a natural experiment, we find that a new train raised demand for elite and more distant schools, but only among high-achieving students with highly-educated parents. These students were also more likely to be assigned to elite and more distant schools under the test-based assignment mechanism. In contrast, we find little effect on the choices or assignments of low-achievers or those with lower-education parents. These results highlight the complementarities between transit access and school choice as well as the potential limitations of choice policies in large urban areas.

JEL Codes: I24, I25, O18, R41
Keywords: mass transit; school choice; inequality

*E-mail addresses: andrew.dustan@vanderbilt.edu, dngo@oxy.edu.

We thank Brady Fuller and Frank Jiawei Ye for invaluable research assistance, Mexico's Secretariat of Public Education for providing the data, and participants at the Western Economics Association International Conference, Occidental College Workshop on School Choice, Pacific Conference for Development Economics, Association for Education Finance and Policy Annual Conference, Midwest International Economic Development Conference, University of Las Vegas, Nevada economics seminar, Italian Development Economist Association Summer School in Development Economics, and Association for Public Policy Analysis and Management International Conference for useful feedback. This research was funded in part by a grant from the Occidental College John Parke Young Initiative on the Global Economy.

1 Introduction

School choice policies are broadly aimed at improving educational outcomes by allowing students to access options that outperform their neighborhood schools or to select schools with attributes that best match their individual needs. Such policies have the potential to increase equity and efficiency and have been introduced in a wide variety of contexts throughout both the developed and developing world.[1] When the geographical distribution of school quality and other attributes is uneven, the success of a choice system requires that students can and will travel farther to attend high-quality, good-match schools (Hastings, Kane, and Staiger 2008). Moreover, similarities or differences in students' ability and willingness to travel will determine whether or not choice policies expand effective access to high-quality schools for all types of students, a debate that is ongoing and relevant to contemporary policy discussions (Hoxby 2003).

In studying the many factors that constrain school access, the literature consistently shows that students highly value proximity.[2] Therefore, distance may continue to limit access to schools that students would otherwise prefer. This is of particular concern for equitable access when distance constrains students differentially with respect to their socioeconomic status (SES). Burgess et al. (2015) suggest that one avenue to increase educational access is "reducing the link between home postcode and the set of schools to which access is feasible in practice" (p. 1288). One potentially attractive avenue for achieving this is the expansion of transit services between where students live and where they might like to attend school, since this may expand access without the difficulties of actively reshaping the geographical distribution of school characteristics.

This paper examines the effects of transit expansion on school choice behavior, using the introduction of the Suburban Train in the Mexico City metropolitan area as a natural experiment. As one of the largest, most populous metropolises in the world, Mexico City offers a unique opportunity to understand choice in an expansive urban area. The Suburban Train was introduced in 2008 and expanded transit access to the relatively poor northern suburbs of the city. It resulted in a substantial change in commuting access to the city center for approximately five million commuters, reducing daily one-way travel times by over fifty percent from 1.5 to two hours before the train to 25 to 45 minutes after (Pantoja and Montaño 2007; Pantoja 2008; Ferrorcarriles Suburbanos 2008a). The new line represented

[1]See Dustan (2017) for an incomplete list of studies covering school choice in various settings. These include studies of choice policies in primary, secondary, and tertiary school systems in Ghana, Kenya, Malawi, Trinidad and Tobago, Romania, China, the United Kingdom, and the United States.

[2]For example, this is a finding in Hastings, Kane, and Staiger (2008), Alderman, Orazem, and Paterno (2001), Gallego and Hernando (2010), and Glick and Sahn (2006), reviewed in more detail in the next section.

lower transportation costs for students, parents, and neighbors in areas surrounding the new line, potentially causing changes in job location, peer information, and community development and crime patterns. This large shock, coupled with a unified choice process and a large number of heterogeneous school options, provides a good context to observe if and how students' choices respond to an expansion in the feasible choice set caused by a new transit line. Notably, the vast majority of Mexico City's "elite" public high schools are situated in the city center and there are none located in the region primarily impacted by the expansion, so students in these areas potentially experienced much greater access to high-quality schools.

We estimate a difference-in-differences model to identify the causal effect of the Suburban Train on school choice behavior and resulting assignment outcomes. We use a rich dataset of the full school rankings for all public high school applicants in Mexico City for each year from 2005 through 2011. The existence of multiple pre-treatment time periods allows us to relax the parallel trends assumption by introducing location-specific linear time trends. The key identifying assumption is that, net of any aggregate temporal shocks, any deviation of treatment area outcomes from their trends in the post-expansion period is due to the existence of the Suburban Train. We identify treatment areas as postal codes close enough to new stations that students may access them and perform the analyses using two sets of control areas. The first draws control areas from the rest of the suburban ring, composed of regions that are comparable distances from the city center and from public transit prior to the Suburban Train. The second uses areas around planned but unbuilt transit lines as control regions.

In our preferred specification, we use the distance reduction to a transit station induced by the train as a measure of treatment intensity, to allow for differential effects between central and remote areas. We use subcluster wild bootstrapped p-values to account for geographical clustering in the presence of a small number of treated clusters, as suggested by MacKinnon and Webb (2017). We also present various robustness checks and show that the results are robust to the empirical specification, varying criteria for identifying the treatment and control areas, and the pre-treatment period used.

We find two main results. First, the train causes some types of students to choose schools that are farther away from home and increases their likelihood of listing elite schools as their first choices. While evidence for the average effect of the expansion on the affected population is weak, we find consistent, statistically significant positive effects concentrated among high-achieving, high-SES students. Following these choice changes, these students are also more likely to be assigned to elite schools and to more distant schools under the exam-based school assignment mechanism.

For a high-scoring, high-SES student experiencing the mean reduction in distance to nearest station (9 kilometers, or a 75% decrease), the implied effects are a 5% increase in distance to first choice, a 2% increase in distance to assigned school, and a 3% to 4% increase in elite first choice and elite assignment. For high-scoring, high-SES students at the end of the Suburban Train (with a reduction in distance to nearest station of 18 kilometers), the implied effects are a 10% to 11% increase in distance to first choice, a 4% to 5% increase in distance to assigned school, and a 7% to 8% increase in elite first choice and elite assignment.

Second, the transit introduction has little effect for other student types, and there is little evidence of substitution toward high-quality nonelite schools. Specifically, we can rule out large effect sizes on distance to chosen schools for low-achieving students and for low-SES students.

Taken together, the findings suggest that the newly constructed stations allowed high-ability students from high-SES backgrounds to access elite high schools that were previously out of reach. Outcomes for other groups remained mostly unchanged either because latent demand for elite schools was low, commuting costs remained too high, or because the train did not dramatically change commute times (in the case of stations near the city core).

Our paper contributes to the literature on school choice policies. While other studies have identified strong preferences for nearby schools in descriptive cross-sectional analyses, we are among the first to identify the causal effects of increased transit access on school choice.[3] This is important because endogeneity issues associated with residential location and school choice suggest that cross-sectional analyses may not accurately predict the effects of policies directed at changing the effective proximity to high-quality schools.

In using this natural experiment, we more narrowly identify the causal effects of constructing new transit stations on school choice. The results suggest that similar policies that cause large changes in transit patterns for students and the broader population can be effective at raising demand for more distant, high-performing schools for certain student groups. Despite the small outcome elasticities with respect to reduction in distance to nearest station, the magnitude of effects on elite choice are comparable to policies directly designed to change students' choices.[4]

However, similar transit policies may not cause a majority of students to substitute away from neighborhood schools toward more distant higher-quality options for a variety of reasons. Students may continue to be constrained by costs, admissions policies and enrollment

[3]Working papers by Herskovic (2017) and Asahi (2015) study the effects of a subway expansion in Santiago, Chile.

[4]Hastings and Weinstein (2008) find that providing families with information about school academic performance causes a 5 to 7 percentage point increase in the fraction of students choosing nonguaranteed schools and a 0.05 to 0.10 standard deviation increase in the average test scores of chosen schools.

constraints, or a lack of information. This suggests that additional contextualized policies may be necessary to achieve effective access for all students. Moreover, low-achieving and low-SES students may have strong preferences for convenience and neighborhood peers and may choose proximate schools even after some constraints are relaxed. In this case, increasing access may require locating high-quality schools close to targeted students rather than moving students to high-quality schools.

The rest of the paper is organized as follows. Section 2 gives a summary of the relevant literature. Section 3 provides detailed overviews of the institutional contexts in Mexico City, describing the geographical setting, the introduction of the Suburban Train, and the system of school choice. Section 4 sets forth the method for identifying the effects of the transit introduction on outcomes. Section 5 describes the data and Section 6 gives the empirical results. Section 7 concludes.

2 Literature review: school choice, willingness to travel, and transit

There is a large body of research describing student preferences for school characteristics. This literature consistently shows that families strongly value proximity in choosing schools. Proximity is an important determinant of choice throughout both the developing and developed world (Hastings, Kane, and Staiger 2008; Burgess et al. 2015; Alderman, Orazem, and Paterno 2001; Chumacero, Gómez, and Paredes 2011). The preference for close schools appears at all grade levels and regardless of whether students are choosing among traditional public schools or among non-traditional options.

Studies also cite transportation convenience as an important determinant of school choice. In Charlotte, North Carolina, Hastings, Kane, and Staiger (2008) find that families are more likely to choose schools within a zone of assured school bus transportation. Outside of the U.S., public transit is particularly important, accounting for a moderate to large mode share of trips to schools (Müller, Tscharaktschiew, and Haase 2008). It is more important in urban settings and in areas where public transit is more available (Schwanen and Mokhtarian 2005). For example, transit access is important in Santiago, Chile, where families are more likely to choose schools that are close to subway stations (Gallego and Hernando 2010). Similarly, in Tel Aviv, Lavy (2006) finds that students who reside along public bus lines leading directly to their district schools are more likely to choose those schools compared to students residing farther from the bus lines.

In addition to proximity, researchers find that families value schools' academic performance in most contexts and that students are willing to travel farther to attend high-performing schools. For example, Hastings, Kane, and Staiger (2008) report that on average, families in Charlotte, North Carolina would be willing to travel an additional 3.8 miles to attend a school with a one standard deviation higher average test score. Studies generally find that the willingness to trade-off between achievement and distance differs systematically by school type and family characteristics. Students are willing to travel farther at the middle and high school levels (Hastings, Kane, and Staiger 2008; Chumacero, Gómez, and Paredes 2011). They also tend to travel farther for private, religious, and charter schools compared to traditional public schools (Butler et al. 2013; Alderman, Orazem, and Paterno 2001; Glick and Sahn 2006). Families with higher incomes and higher levels of parental education are usually willing to travel farther to attend high-performing schools (Hastings, Kane, and Staiger 2008; Chumacero, Gómez, and Paredes 2011; Schneider, Elacqua, and Buckley 2006; Burgess et al. 2015), although Gallego and Hernando (2010) do not detect this relationship. High-achieving students and students with high-achieving peers are also willing to travel farther (Hastings, Kane, and Staiger 2008; Ajayi 2013).

Overall, the school choice literature suggests that increased access to public transportation can affect choice by reducing effective distance, inducing students to substitute toward higher-quality or better-match schools that are farther away.[5] This effect is more likely to be found in urban areas where public transportation is relevant and for students who are more willing to trade off distance for academics or are less constrained by other factors.

In using a transit expansion as a natural experiment, we also build on a growing literature in urban economics that estimates the impact of transportation on various outcomes. Within this literature, the effects of transportation are mostly positive. Various studies find that transit expansions increase job accessibility and employment (Holzer, Quigley, and Raphael 2003; Rotger and Nielsen 2015; Moreno-Monroy and Ramos 2015). At the community level, transit expansions have had positive impacts on land and housing values near newly accessible transit in some cities (Kahn 2007; Billings 2011; Gibbons and Machin 2008), but negative impacts in others (Glaeser, Kahn, and Rappaport 2008). In addition, transit innovations have resulted in changes in local crime (Billings, Leland, and Swindell 2011; Phillips and Sandler 2015) and decreased traffic (Anderson 2014). This literature establishes credible methods for estimating the effects of transit access, which we discuss in more detail below, and also indicates that transportation can have effects on a variety of outcomes. These broad effects also raise the possibility that in addition to reduced commuting time, new

[5] In addition to affecting choice, prior work has found that distance to schools affects educational outcomes. Lower distances to schools are associated with an increased probability of graduating from upper secondary school in Norway (Falch, Lujala, and Strøm 2013) and an increased probability of continuing into post-compulsory education in England (Dickerson and McIntosh 2013).

transit stations could change school demand through other mechanisms, by raising housing values and influencing other neighborhood characteristics.

To our knowledge, there are few studies that examine the effects of transit expansions on school choice outcomes. In Santiago, Chile, Herskovic (2017) finds that the introduction of a new subway line causes affected students to travel to farther schools that perform better on standardized tests. Using the same expansion, Asahi (2015) finds that the introduction of the subway increased enrollment for schools located near new stations. Thus, prior work suggests that changes in school accessibility can have significant effects on school choice, but additional work is necessary to determine if and how this effect manifests in different contexts. Our study complements these two studies by examining a different context with distinct geographical and institutional constraints and by examining heterogeneity by student groups, which is not done in the Santiago studies.

3 Institutional background

3.1 Mexico City transit and the Suburban Train

Mexico City is large and densely populated, and its residents travel extensively throughout the city daily between their homes, jobs, and schools. The Mexico City Metropolitan Area spans 5,300 square kilometers and is comprised of the Federal District (Distrito Federal) and the surrounding State of Mexico (Estado de México). The city is connected through networks of roads, high-capacity public transportation, and medium- and low-capacity informal transportation. Transit ridership is high, with approximately two thirds of all trips relying on public transportation (Guerra 2014), and transit users tend to be poorer than those who drive (Gilat and Sussman 2003). The majority of travel relies on informal private bus routes or on combinations of the Metro, informal transport, and other modes (Guerra 2014; Gilat and Sussman 2003).

The suburban areas in the State of Mexico account for half of the metropolitan population and are less connected to formal public transportation than the city core. Compared to the Federal District, the suburban areas are more densely populated, accounting for the majority of population growth since 1970 (Guerra 2014). They are also poorer and have lower rates of car ownership. Suburban residents utilize the Metro but are more reliant on informal transit, as most Metro lines terminate at the Federal District-State of Mexico boundary. These informal routes are generally costlier and substantially slower than the Metro (Guerra 2014; Gilat and Sussman 2003). In total, prior to the Suburban Train, approximately five million residents in the suburb relevant to our analysis commuted round-trip for three to four hours daily to both jobs and schools in the city center (Pantoja and Montaño 2007; Pantoja 2008).

In recent years, the Mexican government has begun to invest in high-capacity suburban transit in an effort to facilitate commuting and development (Guerra 2014; McKegney 2012). Line 1 of the Suburban Train is among the first of these investments to be completed, running 27 kilometers from the Federal District into the northern suburbs of the State of Mexico.[6] Planning for the Suburban Train was initiated in the 1990s, with a plan including Line 1 as well as two other lines in the eastern suburbs, which were to be started later (Hernandez and Romano 1999). The placement of the Suburban Train was chosen to utilize an existing but unused railroad network, so that the only rail work would involve raising or lowering the tracks in certain places. This existing rail network has a long history, dating back in its present coverage to at least 1976 and taking a very similar path since early in the 20th century (Minsk 2003). This plan was appealing in its large cost savings compared to alternatives such as expanding the Metro, which was projected to cost more per kilometer by a factor of six (Melgar 2003).

While the Suburban Train Line 1's route was thus predetermined by the historical rail network (and indeed, its eventual path conformed to the plan of utilizing this existing infrastructure), the placement of stations along the route was a decision made during the planning and bidding phase (Ministry of Communications and Transportation 2005). The two stations in the Federal District, Buenavista and Fortuna, used existing station infrastructure, while the other five, more distant stations were placed based on the availability of land and proximity to areas with higher population density. These new stations were designed to be centers of economic activity and include pedestrian and bicycle accessibility and designated transfer points to many informal transit lines and to two existing Metro lines in the Federal District (McKegney 2012; Soto 2008).

The period of time between Line 1's announcement, financing, and eventual completion was long and plagued by unexpected delays. Bids for operating companies were solicited in 2003, about four years after plans for the system were made public. The 2003 round of bidding failed, and a second call for bids resulted in the existing contract that was awarded in 2005. This initial contract allowed for a five to ten year period for completion of the system (Construcciones y Auxiliar de Ferrocariles 2006; Secretaria de Comunicaciones y Transportes 2005). The first concrete news for residents regarding the train's operating details came in late 2007, between the 2007 and 2008 school choice cycles, when fares and definitive plans to open in the first half of 2008 were announced in the press (Notimex 2007). Operations began

[6]Metro's Line B, completed in 2000, was a much earlier expansion into the northeastern suburbs. Guerra (2014) provides details.

in May 2008 for the initial five stations (Buenavista, Fortuna, Tlalnepantla, San Rafael, and Lechería) and in January 2009 for the farthest two stations (Tultitlán and Cuautitlán), with the latter delay being unanticipated and due to issues with materials theft and conflicts over right of way (Ortiz 2008; Cruz 2008; Notimex 2008). In total, the project cost approximately $700 million USD.

In 2008, the ride cost 5.5 pesos for a shorter trip (three or fewer stations) and 12.5 pesos for a longer trip (four or more stations), with an additional one-time fee of 11.5 pesos for a rechargeable card (Ferrorcarriles Suburbanos 2008a, 2009).[7] Daily ridership has been increasing from approximately 30,000 passengers in 2008 (Ferrorcarriles Suburbanos 2008b) to 179,000 in 2015 (Miranda 2015). With frequent trains (every six minutes during peak hours and every fifteen minutes during off-peak hours) and extensive hours, the line reduced one-way travel time from an average of two hours to 25 minutes between Buenavista and Lechería (Ferrorcarriles Suburbanos 2008a). Overall, there were large potential impacts for commuters who resided far from the city core.

3.2 School choice in Mexico City

Since 1996, Mexico City has implemented a unified policy for school choice for all students residing in the Mexico City metropolitan area who want to attend a public high school.[8] The admissions process is run by Comisión Metropolitana de Instituciones Públicas de Educación Media Superior (COMIPEMS), a consortium of schools formed to address prior inefficiencies in high school enrollment. Students (as of 2011, the final year examined in this paper) choose from nearly 600 academic programs spread over approximately 360 campuses.[9] Schools are categorized under nine different subsystems, which vary in their curricula and levels of academic rigor. Two of the subsystems are considered elite and are affiliated with prestigious national universities. One elite subsystem is affiliated with the Instituto Politécnico Nacional (IPN) and focuses on sciences and engineering, while the other elite subsystem is affiliated with the Universidad Nacional Autónoma de México (UNAM) and offers a broader curriculum. Both elite subsystems are highly competitive and are highly demanded.[10] The majority

[7]The average exchange rate in May 2008, when the first stations opened, was 10.44 pesos per USD.

[8]Several existing studies in economics have examined various aspects of Mexico City's public high school system. These include Dustan, de Janvry, and Sadoulet (2017) and Estrada and Gignoux, which study the effects of elite schools on various outcomes, and Bobba and Frisancho (2016) and Dustan (2017), which examine behavioral and informational determinants of choice.

[9]Some campuses offer multiple programs, which students may list separately in their ranked options. For our analyses, we treat each program as a unique school. Since all elite schools offer only one program, this does not affect the substance of our analysis.

[10]See Dustan, de Janvry, and Sadoulet (2017) for a more detailed description of the elite schools as well as their associated returns.

of elite schools (28 of 30) are located in the Federal District but within commuting distance of suburban areas. The remaining subsystems consist of a mixture of vocational, technical, and traditional schools that tend to draw students from their local catchment areas. Figure 1 shows the locations of the elite and nonelite schools, as well as the public transit lines that serve the city. To give an example of the pattern of school demand in the northern suburbs, Figure 2 shows the locations of the first choice schools for students at the end of the (not yet opened) Suburban Train line in 2007. While many students chose schools close to home, others chose elite schools in the city center despite the long commutes that these entailed.

The selection process occurs during and after students' final year of middle school (grade nine). In February, students rank their preferred schools, listing up to a maximum of twenty options.[11] In June, students then take a comprehensive placement test.[12] School assignments occur in July, following a serial dictatorship algorithm that is a special case of the deferred acceptance algorithm characterized by Gale and Shapley (1962). This algorithm orders all students according to their placement test scores, treating each student in turn beginning with the highest-performing student. A computer then proceeds to assign each student to her highest-ranked school with a remaining seat when her turn arrives. All students with the same score requesting seats in the same school are accepted or rejected as a group depending on the preference of the school's subsystem, which is consulted in real-time when a tie occurs.[13]

For our analysis, the assignment mechanism has the important property that truthful ranking is a dominant strategy for all families (Dubins and Freedman 1981). Additionally, students have considerable information about the quality and competitiveness of available schools prior to listing their rankings. Official information about individual schools is available to students, with historical admission cutoff test scores posted on the website used for school registration (Dustan 2017). Furthermore, students can attend an annual information fair to speak with representatives from each subsystem. Together, the assignment mechanism, large maximum number of options, and availability of information suggests that students' listed options are reasonable measures for their underlying preferences.

Given this expansive and unified choice system, the substantial change in transit access as-

[11] Only about 3% of students in our sample exhaust their 20 possible choices.

[12] Students who have requested an UNAM-affiliated school as their first choice take a different version of the test, although the two versions of the test are written in collaboration and there does not seem to be a public perception that the UNAM exam differs in difficulty. The reasons for this separation of test versions are historical and political.

[13] About 12% of students who otherwise qualify for assignment are left unassigned by the computerized assignment process because they chose only schools with cutoff scores higher than their own score. These students must later choose from schools that have seats remaining after the computerized assignment, a process for which we do not observe the outcome.

sociated with the Suburban Train had the potential to generate large effects on school choice for the affected population. Moreover, other neighborhood changes associated with new station construction could change school demand, if affected areas also experienced increased safety and housing values. We hypothesize that due to the Suburban Train, students would be more likely to choose more distant schools, particularly the newly commute-accessible elite high schools located in the city center at the end of the line.[14] To the extent that elite preference is correlated with student characteristics, we expect the effects of transit expansion to be strongest among students from more-educated backgrounds and who have higher levels of observed ability.

4 Empirical method

The goal of this paper is to estimate the effect of the Suburban Train's introduction on students' school choice behavior. To do so, we treat the opening in 2008 as an exogenous shock in transit availability to students residing along the new route. Clearly, the location of the train line was not randomly determined, and this placement is potentially correlated not only with students' demographic characteristics but also their school choices. The panel dimension of our data allows us to control for fixed differences in these variables across space. In addition, the placement of the train line was perfectly compliant with the route from the initial plan created in the 1990s, which in turn was fully determined by the location of pre-existing railway infrastructure. Redding and Turner (2015) present a classification of instrumental variables for placement of transportation infrastructure that is useful here: the Suburban Train Line 1 is a case in which, in the cross-section, the actual infrastructure is perfectly predicted by both the "planned route" IV, where the IV is the originally planned Line 1, and the "historical route" IV, where the IV is the location of the existing railway stretching from the Federal District to the boundary of the northwestern suburbs.[15]

While the fact that the Suburban Train conformed to both original plans and historical infrastructure is encouraging, there are two further concerns when treating the expansion as a natural experiment. First, while the route followed by the line was predetermined, the location of the stations was decided during the planning and bidding phase of the project. We will thus adapt the "planned" and "historical route" IVs to the case of station placement,

[14] It is possible that the Suburban Train could have caused students to be more likely to choose other nonelite, commute-accessible schools. We also test for this hypothesis but focus primarily on elite schools, since demand for non-neighborhood schools is driven primarily by demand for elite schools, as seen in Figure 2.

[15] Interpreting this as a historical route IV is most appropriate, since the plan in this case is created by the same set of actors as those implementing the project and the time lag between planning and construction is shorter than the example of the sprawling highway network in Redding and Turner (2015).

instrumenting a student's distance to the nearest station with his distance to the nearest point on the planned (also historical) railway line. This approach, explained in more detail below, yields results almost identical to specifications that do not account for endogenous placement of stations. We therefore present the OLS results as the primary specification and show the IV analysis as a robustness check.

The second concern is that placement and opening of the line is correlated with trends in demographics and choices over time. We address this concern in two ways. The first is econometric, controlling for linear trends in outcomes at the local level. This approach will be explained further below. We will also show that there was no sudden trend break in demographic characteristics of students corresponding to the Suburban Train's opening. The second justification is historical. As we explained in the previous section, there were significant delays between the Train's announcement, the beginning of construction, and the actual opening of the train. There was also substantial uncertainty surrounding the precise date of its opening, which was only resolved in early 2008. Capitalization of land values and resulting demographic shifts that could affect the composition of students are thus unlikely to have occurred suddenly in the time period around 2008. This uncertainty about the precise school cycle in which the Train would start operations is important for thinking of the 2008 opening as a natural experiment. Students in the 2007 cycle and before were unable to count on the existence of the Train as a commuting option, and even lacking this option for one year was likely sufficient to deter many students from choosing schools for which commuting time was (at the margin) a determining factor.

We therefore use the Suburban Train's introduction as a natural experiment, taking the following empirical approach. First, we locate the geographical areas (postal codes) that are likely to be affected by the expansion due to proximity to the Suburban Train stations. Second, we identify similar but unaffected areas using two strategies drawn from the urban economics literature. Finally, we take a linear difference-in-differences style estimation approach that allows for differential trends in outcomes across space, and expand this to an IV robustness analysis that accounts for potentially endogenous station placement. We discuss the selection of treatment and control postal codes in the next section and follow with the estimation strategy.

4.1 Treatment and control postal codes

In order to estimate the change in school choice behavior induced by the Suburban Train, we need to identify areas affected by the expansion and comparable areas to serve as counterfactuals. We describe our preferred treatment and control areas here and present several robustness checks in Section 6.6, where we vary the criteria for inclusion in the treatment and control groups.

Guerra (2014) shows that the Metro Line B, an urban-suburban transit expansion in northeast Mexico City in 2000 and 2001, mostly serves commuters living within five kilometers of the stations.[16] Thus, we define as treated postal codes those whose closest transit station is a Suburban Train station after its introduction and which have a centroid that is within a five kilometer buffer around the station.[17] We assess the effects of the transit expansion using other buffer sizes as well, while maintaining that the five kilometer range is in line with existing literature, is appropriate given the suburban setting where willingness to travel to a transit station may be higher than in a dense urban core, and includes a reasonably large number of postal codes.

To define a control group against which to compare the postal codes near the Suburban Train stations, we employ two strategies for identifying control postal codes, drawn from the urban economics literature. The first strategy defines a "suburban ring" group, drawing contextually similar but unaffected areas as controls, and the second strategy uses areas around planned but unbuilt stations on other planned suburban transit lines to identify control regions.

The first strategy draws from Glaeser, Kahn, and Rappaport (2008), defining control areas as those that experience no change in nearest station after the expansion and restricting our attention to the subset that is contextually similar to our treated areas, as in Gibbons and Machin (2008).[18]

Specifically, we use the following criteria. First, the postal code centroid must not experience any change in distance to closest station following the transit expansion. Second, to draw areas that have similar pre-train access to transit as the treatment areas, the postal code centroid must be at least as far from a transit station as the minimum pre-train distance

[16] The text in Guerra (2014) explains that most commuters are within three kilometers. Figure 5 in Guerra (2014) shows graphically that there are many commuters beyond this three kilometer range, but close to zero beyond five kilometers.

[17] We exclude Buenavista and Fortuna, the last and second-to-last stations in the city center, because they had existing Metro stations prior to the opening of the Suburban Train.

[18] A related strategy compares treated areas close to new transit stations to control areas farther from those same stations (e.g., using a zero to five kilometer radius around a transit station to define treatment areas and a five to ten kilometer radius to define control areas) (Holzer, Quigley, and Raphael 2003; Rotger and Nielsen 2015; Sari 2012; Kahn 2007). We do not employ this strategy since the high willingness to travel in the context of Mexico City suggests that the new transit stations may affect residents who live relatively far from the station (e.g., outside a 5 kilometer radius), but who still experience a decrease in distance to nearest transit station after the expansion. Specifically, we restrict our control areas to those that experience no change in nearest station, while other studies include some control areas that do register a change in nearest station, arguing that those areas remain unaffected by the expansion since they are relatively distant from the new stations.

to station among treatment group postal codes and no farther from a transit station than the maximum pre-train distance to station among treatment group postal codes. Third, we exclude the isolated eastern municipality of Texcoco despite the fact that some of its postal codes meet the inclusion criteria, although we will show that the results are robust to including it.[19] In a separate robustness check, we exclude previously identified control areas that are within the Federal District, in order to assess the possibility that the observed effects of transit expansion are in part due to differential policy changes in the Federal District and State of Mexico at the time of the expansion. The sample forms a "suburban ring" around the core of Mexico City that, prior to the expansion, was poorly served by government-provided mass transit.

Figure 3 shows the treatment and control groups along with the transit network, including the Suburban Train. Compared to the excluded city center, these suburban areas are more similar to the treated sample in terms of demographics and school choice behavior, as we show below. The suburban ring strategy identifies a relatively large control group, ensuring that any results we find are not due to idiosyncratic behavior in one particular control region of Mexico City.

In the second strategy, we define control areas as those with proposed and planned but unbuilt suburban transportation expansions, following Billings (2011) and Moreno-Monroy and Ramos (2015). Specifically, in 2006, the Mexican government announced its approval to expand the Suburban Train system by building Line 2 to the northeastern suburbs and Line 3 to the eastern suburbs (Ferrorcarriles Suburbanos 2006). We therefore define as control postal codes those whose closest transit station is a planned Line 2 or Line 3 station, maintaining a five kilometer buffer around the control stations similar to our definition of treatment areas.

As a complement to the suburban ring control group, this second strategy addresses some remaining concerns about the endogenous placement of the Suburban Train (Line 1) if the line was placed in areas that would have had large changes in commuting behavior, even in the absence of the new train. By using areas around planned transit stations, we limit our control group to other areas that the government identified as suitable for transit expansion, which are more likely to be similar to the treatment areas in terms of transportation needs, likely effects, and expansion feasibility.

[19] This exclusion is due to a policy change in the 2008 COMIPEMS cycle that drastically limited the number of seats available at Texcoco's most-demanded and most-competitive high school, affiliated with the Autonomous University of the State of Mexico (UAEM). The result was a massive shift in demand toward UNAM and IPN high schools for Texcoco residents in 2008 that was unrelated to transit accessibility, making Texcoco an invalid control region. Texcoco residents would have constituted 2.3% of the control group. First choice demand among Texcoco students for the UAEM-affiliated high school fell from 38% in 2007 to 3% in 2008. First choice demand for this school was only 0.2% outside of Texcoco in 2007 and below 0.1% in 2008.

Figure 4 shows the treatment and control groups using the planned station strategy. This control group contains fewer students than the suburban ring control group. It consists of postal codes in the northeastern and southeastern suburbs that are also included in the suburban ring as well as a few areas that were relatively close to transit prior to 2008 and therefore excluded from the suburban ring.

4.2 Difference-in-differences estimation strategy

After defining the treatment and control postal codes, we employ a difference-in-differences estimation specification with differential municipality-by-treatment trends in outcomes.[20] We describe our preferred strategy here and show additional robustness checks in Section 6.6, where we vary the start year, the specification of trends, and the method for performing cluster-robust inference.

The basic specification is as follows. Let y_{ipmt} be a school choice indicator (e.g., distance to first choice school) for student i residing in postal code p within municipality-by-treatment group m taking part in the COMIPEMS admissions process in year t. Let $train_p$ be an indicator for a postal code being (eventually) treated by the Suburban Train and indicate the time period in which the Suburban Train exists with a dummy variable $post_t$. Because we have multiple years of pre-expansion data, it is possible to include linear trends. We include trends for each municipality-by-treatment group m; postal codes are nested within municipality-by-treatment groups. The basic difference-in-differences specification with trends is:

$$y_{ipmt} = \alpha_p + \gamma_t + \delta_m t + \tau \left(train_p \times post_t \right) + X_{ipt}\beta + \varepsilon_{ipmt}, \tag{1}$$

where α_p are postal code fixed effects that allow postal codes to have different baseline levels of the school choice outcome, γ_t are year fixed effects that allow for flexible global trends in outcomes, δ_m are the linear time trend parameters that allow each municipality-by-treatment group's outcomes to evolve differently from the global trend, and X_{ipt} is a vector of student-specific characteristics. The $(train_p \times post_t)$ term is an indicator for whether postal code p has been "treated" by receiving a nearby Suburban Train station. Under the assumption that $E\left[\varepsilon_{ipmt} | \, p, m, t, (train_p \times post_t)\right] = 0$, τ gives the average effect of Suburban Train expansion on students who live in areas with access to it. This exogeneity assumption requires that deviations from time trends (composed of the flexible global trend and linear municipality-by-treatment trend) in affected areas are due to the transit expansion and not due to other, unobserved factors.[21] In other words, even in the presence of different baseline levels and

[20]Some municipalities include both treatment and control postal codes. Our specification allows the trends of these postal codes to differ from each other.

[21]This type of assumption—about deviations from growth paths rather than about parallel trends—is

trends in school choice outcomes stemming from unobserved differences between treatment and control areas, the exogeneity in the timing of the train's opening is sufficient to identify its impacts as long as there were no unobserved area-specific shocks affecting school choice that coincided with the expansion.[22]

Standard errors are clustered at the municipality-by-treatment area level, following Bester, Conley, and Hansen (2011), who suggest addressing spatial correlation using standard errors clustered at a highly-aggregated level. While the sample used in the main analysis includes students residing in 858 postal codes, these are located inside of only 39 municipalities. Here we face two concerns related to inference. One is the modest number of clusters, which may bias standard errors downward. This can be remedied with a wild bootstrapping procedure. This raises the second concern: the relatively low number of treated clusters in the sample, which MacKinnon and Webb (2017) note can result in poor performance of the wild-cluster bootstrap.[23] We use their proposed solution, the subcluster wild bootstrap, in which each repetition resamples residuals at the observation level and then generates a t-statistic using the cluster-robust covariance matrix. They show that this approach tends to outperform the wild cluster bootstrap, sometimes substantially, in settings where there are few treated clusters.[24] We also present wild-cluster bootstrapped p-values in the Appendix.

Our preferred specification expands upon the basic specification to account for the train's differential effects on transit access, with students near the new end-of-line suburban station experiencing very large changes in transit access and students near the city center experiencing relatively small changes.[25] Following Gibbons and Machin (2005), then, we use the change in distance to nearest station as a measure of treatment intensity, maintaining the five kilometer buffer around new stations to identify treated areas. The specification with continuous treatment is:

---

referred to as a "parallel growths" assumption by Mora and Reggio (2012).

[22]We do not take the approach of estimating a non-linear discrete choice model of school choice, although the COMIPEMS structure of the data allows it. As pointed out by Blundell and Dias (2009) and illustrated in Athey and Imbens (2006), even the simplest non-linear difference-in-differences models are difficult to estimate, and it is not straightforward to accommodate important features of our setting: a very large number of possible alternatives (schools) with correlated characteristics, group-specific time trends, and spatial dependence in the disturbance term. We instead choose a reduced-form strategy that focuses on identifying average marginal effects on the outcomes likely to be affected by the transit expansion.

[23]In the most common use of the wild-cluster bootstrap, where the null hypothesis is imposed when bootstrapping the residuals, the procedure tends to under-reject dramatically.

[24]MacKinnon and Webb (2017) show that the procedure tends to under-reject when treatment clusters are relatively small compared to control clusters. In our case, control clusters have about twice as many observations as treated clusters, implying that our tests may be somewhat conservative.

[25]We note, however, that if transit expansion is mostly facilitating commutes into the city center, it is possible that effects are larger or smaller for newly-served areas that are farther from the city. Specifically, residents in remote treated areas continue to have moderate commute times after the introduction of the train and are required to pay a higher fare when traveling more than three stations into the city.

$$y_{ipmt} = \alpha_p + \gamma_t + \delta_m t + \tau \left(train_p \times distance\ change_p \times post_t \right) + X_{ipt}\beta + \varepsilon_{ipmt}, \qquad (2)$$

where $distance\ change_p$ is the change in distance to nearest station before and after the Suburban Train. All other variables are as described above. We also test higher order functional forms but use the linear specification for ease of interpretation and to facilitate statistical inference using the bootstrapping procedure.

Heterogeneous effects of the expansion with respect to student characteristics are estimated using specifications where all control variables are fully interacted with the discrete dimension of heterogeneity (parental education-by-test score cells), along with separate postal code fixed effects, year fixed effects, and municipality-by-treatment linear time trends for each value of the heterogeneity covariate.

To account for the possibility that station locations were correlated with factors that also influence changes in school demand over the sample period, we will perform a robustness check in which the distance reduction to the closest station is instrumented by a measure of distance reduction that treats every point on the historical railroad segment on which the Train was built as a "station." This specification, then, uses only variation in distance to the historical rail line to determine treatment intensity instead of variation in distance to a new station.

5 Data

To estimate our empirical specification, we use a combination of administrative and geospatial data. The student-level data come from the complete COMIPEMS administrative databases for years 2005 through 2011. While data exist for previous years, we begin in 2005 because trends in school choice variables were highly non-linear prior to this year. We provide support for the choice of years in the next section. These data are collected from each student participating in the COMIPEMS high school assignment process. The data include all ranked school choices and the school to which the student was assigned (if any) during the computerized placement process, which we use in constructing the outcomes of interest. Also included are the student's score on the placement exam, which we normalize by subtracting off the year-specific mean and dividing by the year-specific standard deviation among all COMIPEMS takers. Students self-report information about demographics and other individual and household characteristics. For the purposes of this study, we focus on years of parental education, which we construct using the highest level of education attained by either parent. We also use information on birth order and number of siblings.

We limit the sample to students who are in the ninth grade at a middle school located in the metropolitan area served by COMIPEMS at the time of the exam. This excludes middle school graduates who are re-taking the exam, students outside of the COMIPEMS metropolitan area who expect to move there for high school, and adults who are returning to school. This is done in order to focus on the choices of the typical COMIPEMS participant.

The geographic information systems data come from several sources. Mexico's National Institute of Statistics and Geography (INEGI) provides the location of each postal code. Data on Mexico City's transit network, including the location of each Suburban Train station, are also available from INEGI, but we obtained them in a more accessible form from a different online source.[26] Students' home postal codes are available in the COMIPEMS data, allowing us to geographically locate students according to their postal code centroids. The location of each high school was obtained from the official COMIPEMS website.[27] Combining these data files, we calculate the straight-line distance from each student's home postal code centroid to his nearest transit station (Metro, Metrobus, other light rail, or Suburban Train) in the pre- and post-Suburban Train periods, as well as the straight-line distance from the student's home to his chosen schools.

Table 1 shows summary statistics for students living within five kilometers of the (still unfinished) stations in 2007, the final COMIPEMS round prior to the introduction of the Suburban Train, as well as their corresponding suburban ring and planned station control groups.[28] Keeping in mind that our difference-in-differences identification strategy does not rely on control and treatment areas having the same pre-expansion levels of outcomes or covariates, it is nonetheless worthwhile to see how they compare. Beginning with several outcomes of interest relating to school choice and assignment, we see that the treatment groups are quite similar to their corresponding controls.[29] Compared to students in control areas, students near treated stations choose fewer schools (8.9 schools compared to 9.7 and 10.1 schools in the suburban ring and planned station groups, respectively), were less likely to list an elite school as their first choice (50% compared to 64% and 56%), and chose schools

[26] The unofficial Metro website http://mexicometro.org provides the data in Google Earth format, which has been converted to a standard shapefile format by José Gonzalez at http://www.jose-gonzalez.org/mexico-citys-metro-shapefile.

[27] http://opciones.comipems.org.mx.

[28] Note that the suburban ring and the planned stations control groups are overlapping but not nested. Some of the suburban ring control students would have been served by Line 2 or 3, while some students not in the suburban ring control area (because they were not far enough from existing stations) are also in the Line 2 and 3 group because one of the planned stations would have become their closest mass transit node.

[29] The t-tests for the differences are shown in Appendix Table A.1. Due to the large sample size, most of the differences in means are significantly different from zero at the 5% level, even when the magnitudes are very small.

that were farther from home (9.1 kilometers to first choice school compared to 8.6 and 8.8 kilometers). Despite choosing elite first choice schools less frequently, students in treated areas are just as likely or more likely to be admitted to one, with an acceptance rate of 25% compared to 25% and 20%.

Prior to the introduction of the Suburban Train, treatment area students were on average 5.4 kilometers farther from a transit station than students in suburban ring control areas and 5.9 kilometers farther than those in planned stations control areas. The Suburban Train resulted in treated students being, on average, 8.9 kilometers closer to their nearest transit station, while the change for control students was zero, by construction of the sample.

The better placement performance of students in treated areas can be explained by the fact that they score 0.07 to 0.13 standard deviations higher on the entrance exam. Their parents have, on average, about 0.6 to 0.8 additional years of education and they are 3 to 4 percentage points more likely to attend a private middle school.

A small number of nonelite schools were constructed during the sample period, which could affect the outcomes of interest. In particular, students may choose newly-constructed schools near their homes, decreasing the distance to school measures. To account for this possibility, we include as a control variable the number of schools within 2.5 kilometers of the student's postal code centroid. The mean number of nearby schools is smaller in treatment areas (2.5) than the control areas (3.4 and 4.9, respectively).

For another point of comparison, Appendix Table A.1 shows summary statistics for the treatment groups, the control groups (both suburban ring and planned station), and the city center. When compared to the city core, the treatment and control groups look very similar to each other: their students are far less likely to request an elite school as their first choice, choose schools that are much farther away, score lower on the entrance exam, have less-educated parents, and have far less access to public transit.

6 Results

This section presents evidence that the introduction of the Suburban Train changes the choice behavior and resulting assignments of students who live near the new stations. We find positive effects on distance to chosen schools and on the probability of choosing elite schools among students with more highly-educated parents and who have high test scores. A consequence of these students' altered choices is increased assignment to elite and more distant schools.

6.1 Choice of pre-expansion time period

Several years of pre-expansion data are available, which allow us to include municipality-by-treatment linear trends in the regressions. Figure 5 shows the evolution of the main choice outcomes (whether an elite school was the student's first choice, number of schools chosen, distance to first choice school, and mean distance to first three choices) in the period between 1998 and 2011 for treatment and control (suburban ring) students. There are two salient features of this set of graphs. First, the trends are highly non-linear, with large declines in each measure between 1999 and 2000 followed by recoveries that taper off in 2005. The declines are associated with a massive strike in the UNAM system in 1999 and 2000, which caused demand for UNAM schools to plunge and then recover in subsequent years, with consequences for other subsystems as students substituted toward them. Second, prior to 2005, these trends were somewhat different between treatment and suburban ring control areas. In addition to differential effects of the UNAM strike, the Metro Line B introduction in 1999 and 2000 was also likely to affect demand in one of our control areas, the northeastern suburbs of the metropolitan area. Thus the graphical evidence suggests that including the full pre-expansion period in a model with linear municipality-by-treatment time trends, even in a model with year fixed effects, is likely to be inappropriate.

The change in trends in 2005 suggested by the graphs is confirmed by a simple regression-based comparison of fits in the pre-expansion period between 2001 and 2007, during which the outcomes were generally increasing. We estimate separately for the suburban ring control and treatment groups, for each of the four main outcomes, the following model:

$$y_{ipmt} = \alpha_p + \delta_m t + \beta_1 \mathbb{1}\left(t \geq t^*\right) + \beta_{2m} \mathbb{1}\left(t \geq t^*\right) t + \varepsilon_{ipmt}, \tag{3}$$

where we vary the timing of the trend break (t^*) between 2003 and 2006 and also estimate a model without any break ($\beta_1 = \beta_2 = 0$). For each outcome and group (treatment or suburban ring control), we compare the Akaike Information Criterion between the candidate years for the trend break. In four out of the eight comparisons (two each in the treatment and control samples), the model with a break in 2005 has the minimum AIC, while in the other four, the AIC is minimized in the 2006 break model.[30] Given this and the visual evidence, we choose 2005-2007 as the pre-expansion period for the analysis, favoring 2005 over 2006 as the starting point since it allows us to include three years of pre-expansion data instead of two. Our findings are robust to changing the pre-expansion period, as shown in Appendix Tables A.17 and A.18.

[30]Appendix Table A.2 presents the AIC comparisons.

6.2 Null effects on demographic changes

Before looking at the school choice outcomes of interest, we show that the introduction of the Suburban Train did not correspond to differential changes in the demographic characteristics of students taking the COMIPEMS exam. Table 2 shows the distance change specification for the suburban ring and planned stations control groups in Panels A and B, respectively, using demographic characteristics and exam score as the dependent variables. All results are statistically insignificant, indicating that the observable student characteristics did not change differentially between the treatment and control groups at the time of transit expansion.[31] Thus, within the short time period and for the characteristics we observe, we do not find evidence of students moving due to the transit expansion, indicating that the effects we find below are due to changes in choice among similar types of students.

6.3 Average effects of transit expansion

Turning to school choice outcomes, we present the estimated average effects of transit expansion in Table 3. Columns 1 through 4 present the results of specification 1, the average effect for all treated areas, and columns 5 through 8 present the results of specification 2, using the distance change as a measure of treatment intensity. *Close to station* corresponds to $train_p$ from Equations 1 and 2. We include individual demographics (male, years of parental education, normalized COMIPEMS score, indicators for being the oldest sibling or only child, and number of siblings) and the number of schools within 2.5 kilometers of the postal code centroid in these and all following regressions.

In the binary treatment specification (columns 1 through 4), we find no consistent evidence that, on average, the Suburban Train induces students to change their choice behavior in the period between 2008 and 2011. The point estimates in each case are zero or positive (reporting here the results from the suburban ring control group), representing a 0.4 percentage point increase in the probability of choosing an elite school as first choice, 0.14 additional schools chosen, a 0.11 kilometer increase in distance to first choice school, and an average 0.16 kilometer increase in distance to first three choices. Only the distance to first three choices is (marginally) statistically significant. The point estimates from regressions using the planned station control group are similar but none of the estimated effects are statistically significant.

Columns 5 through 8 examine the effects using treatment intensity (the change in distance to nearest station) measured in 10 kilometer units so that the coefficients approximate the implied effects for the mean treated student, who experiences an 8.9 kilometer distance

[31] The specifications with a binary indicator for treatment show similar results.

reduction. There are marginally significant effects on the distance to first choice and distance to first three choice outcomes. The coefficients in Panel A, using the suburban ring control group, indicate that students who experience a 10 kilometer decrease in distance to nearest station choose schools that are 0.17 kilometers farther (first choice) and 0.19 kilometers farther (first three choices). The corresponding changes for the mean treated student are 0.15 and 0.17 kilometers, respectively, and are 0.32 and 0.34 kilometers, respectively, for students at the most remote end-of-line station. The planned stations control group (Panel B) yields very similar results.

The remainder of the results use the distance change specification, since this provides a more nuanced understanding of differential effects with respect to treatment intensity.

6.4 Heterogeneous effects of transit expansion

We now examine heterogeneous effects with respect to two dimensions of student characteristics: parental education (to proxy for SES) and student achievement. We partition the sample into high and low parental education, where high parental education (which we use interchangeably with high-SES) refers to completion of high school or beyond; this categorization splits the sample roughly in half. We use students' scores on the entrance exam as measures of achievement, partitioning the sample into high and low achievement using above versus below the year-specific median exam score.

Table 4 shows separate effects with respect to parental education-by-student achievement cells. The results show that the positive effects are driven entirely by high-scoring, high-SES students, where the effects are significant at the one percent level for elite first choice, distance to first choice, and distance to first three choices. For this group, the proportion of students listing an elite first choice increases by 2.9 percentage points for each 10 kilometers in distance change compared to a within-group mean (treatment and control) of 78.7% in 2007. For each 10 kilometers in distance change, the distance to first choice school rises by 0.56 kilometers (compared to a 9.69 kilometers base) and the average distance to their first three choices increases by 0.49 kilometers (again compared to a 10.46 kilometers base). For high-scoring, high-SES students at the end of the line, the implied effects are a 5.3 percentage point increase in elite first choice, a 1.02 kilometer increase in distance to first choice, and a 0.91 kilometer increase in distance to first three choices. The planned stations control results are similar in magnitude and significance.

To further illustrate the effects for the high-scoring, high-SES group, Figure 6 presents the results of a simple event study. Each regression includes postal code fixed effects, COMIPEMS exam year fixed effects, student covariates, and treatment-by-year interaction terms. The interaction term for 2007 is excluded so that the treatment-by-year coefficients plotted in the figure represent yearly binary treatment effects relative to the final pre-Suburban Train COMIPEMS cycle. The sample is limited to high-scoring, high-SES students, including the suburban ring control group and treated students with above-median treatment intensity (9.3 kilometer or more reduction in distance to closest transit station). The figure has three notable features. First, while the empirical approach taken in this paper does not rely on parallel trends, we see that the pre-treatment trends for elite preference and distance to first choices are very flat. Second, the treatment effect on these outcomes is clearly visible, beginning with the jump in outcomes in 2008. Third, it seems that these effects increase somewhat over time. This could be due to several factors, such as transmission of school preferences through peer networks or higher levels of certainty about the feasibility of using the Suburban Train as part of a regular commute.[32]

Regardless of the control group selected, we are unable to reject the null of no effect on elite preference and distance to first or first three choices for the other groups. Thus the evidence suggests that, for school choice behavior, transit expansion was most important for students from higher-SES backgrounds and who were strong enough students to expect that elite school admission was a possibility (and a desirable outcome) for them.

Appendix Table A.3 shows that there was little discernible pattern of effects on students' choice of nonelite schools, including the number of options chosen, cutoff score of the most-preferred option, or distance from home. While there is some evidence that high-SES, low-achieving students list more nonelite choices, we do not find that students substitute toward high-quality nonelite schools as shown by the null effect on nonelite cutoff scores. Thus it seems that the main effect of transit expansion was to induce more students to choose elite schools, which are more distant than the nonelite schools chosen by most students.

To explore whether or not the effects are driven by commutes directed into the city, we analyze a set of outcomes related to commute accessibility and distance to the city center. Specifically, we examine schools' proximity to the city center, where the city center is defined by the geometric average of all nodes in Mexico City's subway system. We also look at commute accessibility, where we define "commute schools" as those that are within 2.5 kilometers of any transit station and are not the students' home stations, where home stations are defined as the closest station to the student's postal code centroid. The suburban ring control group results in Table 5 show that for every 10 kilometers in distance change, high-scoring, high-SES students opt to travel closer to the city center by 0.43 kilometers (first choice) and 0.31 kilometers (first three choices) compared to a base of 12.41 kilometers and 12.40 kilometers, respectively. The estimates for the number of commute schools chosen and

[32]The event study for the full sample is in Appendix Figure A.1.

an indicator for choosing a commute accessible first choice are positive but not significant for the high-scoring, high-SES group. For the other three subgroups, the coefficients are mixed in sign and show no clear pattern. Results are similar using the planned station control group and are shown in Appendix Table A.4. Together, the evidence suggests that the main effect of transit expansion was to induce some types of students to choose schools in the city center. Since elite schools are concentrated in the city center and are nearly all close to transit stations, we are unable to separately decompose increased demand for central, commute accessible schools from increased demand for elite schools.

6.5 Effects on student assignment outcomes

The changes in student choice affect the resulting school assignments, as reported in Table 6, which shows the effects by parental education and student achievement for the suburban ring control group.[33] The transit expansion causes high-scoring, high-SES students who were brought 10 kilometers closer to transit to be assigned to schools that are 0.24 kilometers farther away and have a 0.06 standard deviation higher cutoff score, where 2007 cutoff scores are used to capture school competitiveness prior to any potential changes induced by the transit expansion.[34] They are also 2.0 percentage points more likely to be assigned to an elite school. For high-scoring, high-SES students at the end of the line, the implied effects are a 3.7 percentage point increase in elite assignment, a 0.11 standard deviation increase in assigned school cutoff score, and a 0.44 kilometer increase in distance to assigned school. These changes are consistent with the high-scoring, high-SES group's increased choice for elite schools and the fact that they are more likely to actually score high enough for admission to one of their listed elite schools. We find no effects on assigned school outcomes for low-scoring students or for high-scoring, low-SES students, consistent with the null effects on school choices for these groups. The results using the planned station control group are similar and are shown in Appendix Table A.6. The 2 percentage point effect (4 percentage points for end-of-line students) on this group's probability of elite school admission represents a 4% (7% for end-of-line students) increase over the 52% base in 2007. Taken together, the results indicate that the effects on school choice translated to significant changes in assignment outcomes for students who scored high enough for elite admission.

[33] Appendix Table A.5 shows average effects on school assignment; on average, there are positive and significant effects on assigned school cutoff score and marginally significant positive effects on elite assignment.

[34] Students left unassigned by the mechanism are dropped from all regressions in this table. In results not reported here, we find precise null effects of the expansion on the probability of assignment, suggesting that censoring of unassigned students is not driving the reported effects.

6.6 Robustness of results

This section summarizes the results of several robustness checks for the effects of transit expansion.[35]

First, we present results of the IV specification in which distance reduction to the nearest station is instrumented by the reduction measure using distance to the nearest point on the historical rail route used by the Suburban Train. Table 7 presents the results for the suburban ring control group. Appendix Table A.7 presents the first stage results, and Appendix Table A.8 presents the results for the planned station control group. The estimated effects are very similar in magnitude and significance.

To understand whether the estimated effects are sensitive to our choice of a five kilometer buffer for defining postal codes as "treated" by new stations, we estimate effects on the main choice outcomes using three and seven kilometer buffer sizes. The estimated effects for the three kilometer buffer, found in Appendix Table A.9, are larger than those for the five kilometer sample and continue to be statistically significant. For example, the treatment intensity effect on distance to first choice increases from 0.56 to 0.68 kilometers for high-scoring, high-SES students in the suburban ring control group. This is consistent with effects being larger for students in closer proximity to transit stations. Estimated effects for the seven kilometer buffer, found in Appendix Table A.10, are smaller than those for the five kilometer buffer, but they remain statistically significant.

We present results using the criteria specified above but including control postal codes in the isolated Texcoco municipality in Appendix Table A.11. The inclusion of Texcoco only affects the suburban ring control group, since no Texcoco postal codes are within five kilometers of a planned station. The estimated transit expansion effects decline only slightly for the high-scoring, high-SES group and remain statistically significant. Excluding the Federal District from the sample in Appendix Table A.12, we find no evidence that the estimated effects are driven by a State of Mexico- or Federal District-specific shock (for example, a policy change specific to one of these areas in 2008 or later).

With respect to model specification, Appendix Table A.13 presents estimated choice effects from a model including postal code linear time trends instead of municipality-by-treatment linear time trends, while maintaining all other fixed effects and covariates. The results are slightly smaller in magnitude but remain significant. Appendix Table A.14 presents results for a model with no linear time trends. Again, the results for the high-scoring, high-SES group are similar, slightly larger in magnitude and statistically significant. Appendix

[35] We do not calculate bootstrapped p-values for the majority of our robustness analyses, since the significance levels for the results of interest in the main analyses are similar to those using the non-bootstrapped clustered standard errors and the bootstrapping procedure is computationally intensive.

Table A.15 presents the identical specification shown in Table 4 using wild-cluster bootstrapped p-values which, as discussed in Section 4.2, can under-reject the null when there are few treated clusters. While these p-values are more conservative than our preferred wild bootstrapped p-values, the effects for the high-scoring, high-SES groups remain significant at the 5 percent level, with the exception of distance to first choice when using the planned station control group ($p = 0.06$).

A placebo test using pre-Suburban Train time periods fails to find effects of a simulated expansion, strengthening the case that the significant effects are not spurious. The results in Appendix Table A.16 are from a specification using years 2005 and 2006 as the pre-period and 2007 as the post-period. The estimated coefficients are generally small, negative, and quite imprecise. Of the 16 estimated placebo coefficients, only one is marginally statistically significant at the 10 percent level (number of schools chosen for low-SES, low-scoring students).

Finally, Appendix Tables A.17 and A.18 show that the results are not driven by the specific pre-expansion time period that we choose. We find similar effects beginning the pre-period in 2004, 2005 (our preferred specification), and 2006.

7 Discussion

This paper uses the introduction of Mexico City's Suburban Train to identify the effects of transit expansion on students' high school choices and resulting assignments. The transit introduction causes high-achieving, high-SES students located far from the city core to choose more distant schools and increases the likelihood that they choose elite high schools. This results in a higher probability that these students are assigned to elite schools and a higher average distance traveled to their assigned schools. Expansion has no detectable effect on the characteristics of nonelite school choices or the choices of low-achieving or low-SES students.

The analysis has a few limitations. First, we do not observe students' transit mode, so we cannot directly identify those students who actually use the Suburban Train. We are therefore left to speculate about how actual transit usage patterns changed and how these related to changes in school choices.

Second, we cannot decompose the causal mechanisms driving the results. While the Suburban Train decreased commute times substantially, the new transit stations also had the potential to significantly change neighborhood characteristics like home values and safety. Moreover, the train likely changed parental commuting and work opportunities for high school students and graduates, providing increased access to jobs in the city center or at the new Suburban Train stations. These could all contribute to the increased demand for elite schools.

Third, the locations of student residences are approximated by the postal code centroid, introducing measurement error in the distance to station variable that is systematically larger for larger postal codes, which are common in the suburban areas. However, the median area of COMIPEMS-participating postal codes is 0.52 square kilometers, implying that the postal code centroids are likely to be sufficient approximations for student residences.

Fourth, this represents a single case study, and the applicability of these findings in other contexts remains to be seen. Specifically, Mexico City's choice system is unified and transparent. There would likely be smaller effects associated with relaxing commuting constraints in more fragmented choice systems where lack of information and disjointed application policies represent larger barriers. Also, the system we study focuses specifically on high school students, and descriptive research suggests that the impacts on younger students would be smaller (Hastings, Kane, and Staiger 2008; Chumacero, Gómez, and Paredes 2011). On the other hand, the pecuniary costs of commuting are relatively large in our context, with round-trip prices between \$1 to \$2.50 USD relative to an average urban teen wage of about \$2 USD per hour.[36] Free transportation options may induce larger effects or cause low-SES students to also change their behavior. Finally, the importance of transit access is highly dependent on the idiosyncrasies of geographic regions. While the Suburban Train represents a large change in commute time, the overall distance to the city center may still be too far for many students to travel. Similarly, the train's effect is limited to its specific route, with the most direct impact on travel to schools along the train's line and on travel to the elite schools clustered in the city core. A transit opening in a more central location might expand access to a more dispersed set of schools.

Despite these caveats, the findings are instructive and relevant to policies in other contexts. Public transportation captures a large share of school commutes in many urban areas (Müller, Tscharaktschiew, and Haase 2008; Van Ristell et al. 2013) and is particularly relevant in the rapidly growing megacities of many developing countries, which often struggle with traffic congestion and the associated effects of lost time and productivity (Wöhrnschimmel et al. 2008). Despite the long commute times, there is often a high willingness to travel to school (Chumacero, Gómez, and Paredes 2011). In terms of policy relevance, our analysis focuses on a poor suburban area near an urban center. This population also reflects some of the challenges of metropolitan and suburban regions in the United States and other developed countries, which have recently seen increases in suburban poverty that have outpaced increases in urban poverty (Kneebone and Berube 2013; Hunter 2014).

[36]We note that willingness to travel was high prior to the train, and pre-existing modes of informal transportation were either comparable or more expensive in terms of monetary costs.

For these contexts, the results suggest that new public transit lines can expand student choice. This is consistent with the findings by Herskovic (2017) and Asahi (2015), who show large impacts from the new subway in Santiago, Chile, which spans both the city center and a peripheral area. Our results show that public transportation also has effects in areas that are quite far from the city. Reducing very long commutes to moderate commutes is sufficient to change choice behaviors for some types of students. The analysis suggests that treatment intensity is relevant, suggesting that there may be relatively larger effects in areas with lower prior commute accessibility.

With respect to other transportation-related policies, the results demonstrate that some students are willing to commute quite far to high-quality schools. This is one prerequisite for the effectiveness of other transportation-related policies such as introducing school bus lines or providing public transit subsidies to students. However, these other policies may have larger or smaller effects depending on a variety of factors. For example, school bus lines may have larger effects or may affect more types of students if the buses are free or are perceived as safer than public transit. They may have smaller effects if broader population changes are important. With school buses, commuting times may remain long, and job locations and neighborhood characteristics may remain unchanged. Thus, the results raise the possibility that other policies that reduce commuting constraints can be effective, but additional work is required to more carefully identify the important constraints in each context.

Our study of Mexico City adds a different institutional context that is relevant for other achievement-based school choice systems. In Chile's voucher-based school choice system, in which schools set their own admission policies, most students apply to only one school and only 4% are rejected by a school to which they apply (Gallego and Hernando 2010). In contrast, the competitive admissions process in Mexico City may be one factor in generating the large differential effects we find. Low-achieving students and low-SES students may not substitute toward high-quality schools despite relaxed commuting constraints because they may believe that their chances of getting in are low. This is relevant for other merit-based, choice-based assignment mechanisms such as those in Boston, New York, Ghana, Trinidad and Tobago, and Kenya.[37]

Our differential effects ultimately suggest that policies expanding transportation access can be effective for some groups but insufficient for inducing change in others. Thus, careful consideration regarding the location of high-quality schools, constraints for low-achieving students and low-SES students, and admissions processes and enrollment policies remains a central aspect of achieving the various goals of school choice.

[37]See Dustan (2017) for a review.

References

Ajayi, Kehinde F. 2013. "School choice and educational mobility: Lessons from secondary school applications in Ghana." *IED Working Paper* 259.

Alderman, Harold, Peter F Orazem, and Elizabeth M Paterno. 2001. "School quality, school cost, and the public/private school choices of low-income households in Pakistan." *Journal of Human Resources* 36 (2): 304–326.

Anderson, Michael L. 2014. "Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion." *The American Economic Review* 104 (9): 2763–2796.

Asahi, Kenzo. 2015. "Impacts of better transport accessibility: evidence from Chile." PhD thesis, London School of Economics and Political Science (LSE).

Athey, Susan, and Guido W Imbens. 2006. "Identification and inference in nonlinear difference-in-differences models." *Econometrica* 74 (2): 431–497.

Bester, C Alan, Timothy G Conley, and Christian B Hansen. 2011. "Inference with dependent data using cluster covariance estimators." *Journal of Econometrics* 165 (2): 137–151.

Billings, Stephen B. 2011. "Estimating the value of a new transit option." *Regional Science and Urban Economics* 41 (6): 525–536.

Billings, Stephen B, Suzanne Leland, and David Swindell. 2011. "The effects of the announcement and opening of light rail transit stations on neighborhood crime." *Journal of Urban Affairs* 33 (5): 549–566.

Blundell, Richard, and Monica Costa Dias. 2009. "Alternative approaches to evaluation in empirical microeconomics." *Journal of Human Resources* 44 (3): 565–640.

Bobba, Matteo, and Verónica Frisancho. 2016. "Learning About Oneself:The Effects of Signaling Academic Ability on School Choice." *Working paper.*

Burgess, Simon, et al. 2015. "What parents want: School preferences and school choice." *The Economic Journal* 125 (587): 1262–1289.

Butler, JS, et al. 2013. "Choice in a world of new school types." *Journal of Policy Analysis and Management* 32 (4): 785–806.

Chumacero, Rómulo A, Daniel Gómez, and Ricardo D Paredes. 2011. "I would walk 500 miles (if it paid): Vouchers and school choice in Chile." *Economics of Education Review* 30 (5): 1103–1114.

Construcciones y Auxiliar de Ferrocariles. 2006. "The Valle de Mexico Suburban Rail System (ZMVM)." Accessed March 4, 2017. https://web.archive.org/web/20060420114736/http://www.caf.net:80/download.php?file=img/pdf/mexicoingles.pdf.

Cruz, Lilián. 2008. "Sufre retrasos el Suburbano (Suburban Train experiences delays)." Accessed via ProQuest, April 2, 2017, *Reforma, México City.*

Dickerson, Andy, and Steven McIntosh. 2013. "The Impact of Distance to Nearest Education Institution on the Post-compulsory Education Participation Decision." *Urban Studies* 50 (4): 742–758.

Dubins, Lester E, and David A Freedman. 1981. "Machiavelli and the Gale-Shapley algorithm." *The American Mathematical Monthly* 88 (7): 485–494.

Dustan, Andrew. 2017. "Family Networks and School Choice." *Working paper.*

Dustan, Andrew, Alain de Janvry, and Elisabeth Sadoulet. 2017. "Flourish or Fail? The Risky Reward of Elite High School Admission in Mexico City." *Journal of Human Resources* 52 (3): 756–799.

Estrada, Ricardo, and Jérémie Gignoux. "Benefits to elite schools and the formation of expected returns to education: Evidence from Mexico City." *European Economic Review* 95:168–194.

Falch, Torberg, Päivi Lujala, and Bjarne Strøm. 2013. "Geographical constraints and educational attainment." *Regional Science and Urban Economics* 43 (1): 164–176.

Ferrorcarriles Suburbanos. 2006. "Authorizing bids for Suburban Train System 3." Accessed March 4, 2017, *Ferrocarriles Suburbanos, Press Information.* https://web.archive.org/web/20070928220741/http://www.fsuburbanos.com/noticias.php?id=22.

— . 2008a. "Suburban train starts commercial operation." Accessed June 4, 2016, *Ferrocarriles Suburbanos, Press Information.* http://www.fsuburbanos.com/secciones/sala_prensa/noticias.php?id=48&page=2&anio=2008.

— . 2008b. "The Suburban exceeds one million users." Accessed June 4, 2016, *Ferrocarriles Suburbanos, Press Information.* http://www.fsuburbanos.com/secciones/sala_prensa/noticias.php?id=56&page=1&anio=.

— . 2009. "Tultitlán and Cuautitlán opening January 5, 2009." Accessed June 4, 2016, *Ferrocarriles Suburbanos, Press Information.* http://www.fsuburbanos.com/secciones/sala_prensa/noticias.php?id=64&page=2&anio=2009.

Gale, David, and Lloyd S Shapley. 1962. "College admissions and the stability of marriage." *The American Mathematical Monthly* 69 (1): 9–15.

Gallego, Francisco A, and Andrés Hernando. 2010. "School choice in Chile: Looking at the demand side." *Pontificia Universidad Catolica de Chile Documento de Trabajo*, no. 356.

Gibbons, Stephen, and Stephen Machin. 2005. "Valuing rail access using transport innovations." *Journal of Urban Economics* 57 (1): 148–169.

— . 2008. "Valuing school quality, better transport, and lower crime: evidence from house prices." *Oxford Review of Economic Policy* 24 (1): 99–119.

Gilat, Michael, and Joseph Sussman. 2003. "Coordinated transportation and land use planning in the developing world: Case of Mexico City." *Transportation Research Record: Journal of the Transportation Research Board*, no. 1859:102–109.

Glaeser, Edward L, Matthew E Kahn, and Jordan Rappaport. 2008. "Why do the poor live in cities? The role of public transportation." *Journal of Urban Economics* 63 (1): 1–24.

Glick, Peter, and David E Sahn. 2006. "The demand for primary schooling in Madagascar: Price, quality, and the choice between public and private providers." *Journal of Development Economics* 79 (1): 118–145.

Guerra, Erick. 2014. "Mexico City's suburban land use and transit connection: The effects of the Line B Metro expansion." *Transport Policy* 32:105–114.

Hastings, Justine, Thomas Kane, and Douglas Staiger. 2008. "Heterogeneous preferences and the efficacy of public school choice." *NBER Working Paper* 2145.

Hastings, Justine S, and Jeffrey M Weinstein. 2008. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *The Quarterly Journal of Economics* 123 (4): 1373–1414.

Hernandez, Erika, and Luz Romano. 1999. "Proponen ferrocarril suburbano : Aprovecharan vias para transporte masivo (Suburban Train proposal: Taking advantage of routes for mass transport)." Accessed via ProQuest, November 17, 2017, *Reforma, México City.*

Herskovic, Luis. 2017. "The Effect of Subway Access on School Choice." *Working paper.*

Holzer, Harry J, John M Quigley, and Steven Raphael. 2003. "Public transit and the spatial distribution of minority employment: Evidence from a natural experiment." *Journal of Policy Analysis and Management* 22 (3): 415–441.

Hoxby, Caroline Minter. 2003. *The economics of school choice.* University of Chicago Press.

Hunter, Paul. 2014. *Poverty in suburbia: a Smith Institute study into the growth of poverty in the suburbs of England and Wales.* Accessed August 4, 2016. https://smithinstitutethinktank.files.wordpress.com/2014/09/poverty-in-suburbia.pdf.

Kahn, Matthew E. 2007. "Gentrification Trends in New Transit-Oriented Communities: Evidence from 14 Cities That Expanded and Built Rail Transit Systems." *Real Estate Economics* 35 (2): 155–182.

Kneebone, Elizabeth, and Alan Berube. 2013. *Confronting suburban poverty in America.* 10 Facts about Suburan Poverty. Accessed August 4, 2016. Brookings Institution Press. http://confrontingsuburbanpoverty.org/wp-content/uploads/2013/05/Brookings_ToolKit_10-Facts.pdf.

Lavy, Victor. 2006. "From forced busing to free choice in public schools: quasi-experimental evidence of individual and general effects." *National Bureau of Economic Research* 11969.

MacKinnon, James G, and Matthew D Webb. 2017. "The subcluster wild bootstrap for few (treated) clusters." *Queen's Economics Department Working Paper* 1364.

McKegney, Tony. 2012. "Subir Tren Suburbano! Commuter Rail Emerges in Mexico's Largest City." Accessed June 4, 2016, *www.railmagazine.org.* http://web1.ctaa.org/webmodules/webarticles/articlefiles/RAIL_30_Suburbano.pdf.

Melgar, Ivonne. 2003. "Arrancan proyecto del tren suburbano (Suburban Train project starts)." Accessed via ProQuest, November 17, 2017, *Reforma, México City.*

Ministry of Communications and Transportation. 2005. "Título de Concesión para la prestación del servicio público de transporte ferroviario de pasajeros en la modalidad de regular suburbano en la ruta Cuautitlán-Buenavista (Concession for the provision of the Suburban Train)." Accessed November 21, 2017, *Diario Oficial de La Federacion.* http://www.dof.gob.mx/nota_detalle.php?codigo=2095985&fecha=25/10/2005.

Minsk, Todd. 2003. "Research in Mexican railroad history: a chronology of railroad development." Accessed: 11-21-2017. http://www.dartmouth.edu/~rtminsk/queretarotracks1.html.

Miranda, Alan. 2015. "Suburban expects to move 52 million." Accessed June 4, 2016, *Terra Mexico.* http://noticias.terra.com.mx/mexico/estados/espera-suburbano-mover-a-52-millones,0e955a6233bc2976dc0075866eb5b722eltuRCRD.html.

Mora, Ricardo, and Iliana Reggio. 2012. "Treatment effect identification using alternative parallel assumptions." *Universidad Carlos III de Madrid Working Paper* 12-33.

Moreno-Monroy, Ana Isabel, and Frederico Ramos. 2015. “The impact of public transport expansions on informality: the case of the São Paulo Metropolitan Region.” *Working paper.*

Müller, Sven, Stefan Tscharaktschiew, and Knut Haase. 2008. “Travel-to-school mode choice modelling and patterns of school choice in urban areas.” *Journal of Transport Geography* 16 (5): 342–357.

Notimex. 2007. “Aplicará dos tarifas el Tren Suburbano. (Suburban Train will apply two rates).” Accessed via ProQuest, April 2, 2017, *Notimex, México City.*

— . 2008. “Pendientes obras complementarias del Tren Suburbano en Tultitlán (Additional work on the Suburban Train in Tultitlán).” Accessed via ProQuest, April 2, 2017, *Notimex, México City.*

Ortiz, Francisco. 2008. “Desaceleran al Suburbano (Suburban Train delay).” Accessed via ProQuest, April 2, 2017, *Reforma, México City.*

Pantoja, Sara. 2008. “With the Suburban Train, users will rest more.” Accessed June 4, 2016, *El Universal.*

Pantoja, Sara, and Teresa Montaño. 2007. “The Suburban Train seeks to relieve the ordeal of commuters.” Accessed June 4, 2016, *El Universal.*

Phillips, David C, and Danielle Sandler. 2015. “Does public transit spread crime? Evidence from temporary rail station closures.” *Regional Science and Urban Economics* 52:13–26.

Redding, Stephen J, and Matthew A Turner. 2015. “Transportation costs and the spatial organization of economic activity.” In *Handbook of regional and urban economics*, 5:1339–1398. Elsevier.

Rotger, Gabriel Pons, and Thomas Sick Nielsen. 2015. “Effects of Job Accessibility Improved by Public Transport System: Natural Experimental Evidence from the Copenhagen Metro.” *European Journal of Transport and Infrastructure Research* 15 (4): 419–441.

Sari, Florent. 2012. “The Effects of the Tramway on Mobility and Unemployment in Bordeaux: A Difference-in-Differences Analysis.”

Schneider, Mark, Gregory Elacqua, and Jack Buckley. 2006. “School choice in Chile: Is it class or the classroom?” *Journal of Policy Analysis and Management* 25 (3): 577–601.

Schwanen, Tim, and Patricia L Mokhtarian. 2005. “What affects commute mode choice: neighborhood physical structure or preferences toward neighborhoods?” *Journal of Transport Geography* 13 (1): 83–99.

Secretaria de Comunicaciones y Transportes. 2005. "Labor Report Five 2004-2005." Accessed March 4, 2017. http://www.gob.mx/cms/uploads/attachment/file/67738/Informe_de_labores_2005.pdf.

Soto, Mariana. 2008. "Suburban Train is ready, begins operation in March." Accessed June 4, 2016, *Milenio*. http://www.fsuburbanos.com/uploads/prensa/pdf/2008-01-02-MILENIO.pdf.

Van Ristell, Jessica, et al. 2013. "Quantifying the transport-related impacts of parental school choice in England." *Transportation* 40 (1): 69–90.

Wöhrnschimmel, Henry, et al. 2008. "The impact of a Bus Rapid Transit system on commuters' exposure to Benzene, CO, PM 2.5 and PM 10 in Mexico City." *Atmospheric Environment* 42 (35): 8194–8203.

Figures

Figure 1: Locations of schools and public transportation (2011)

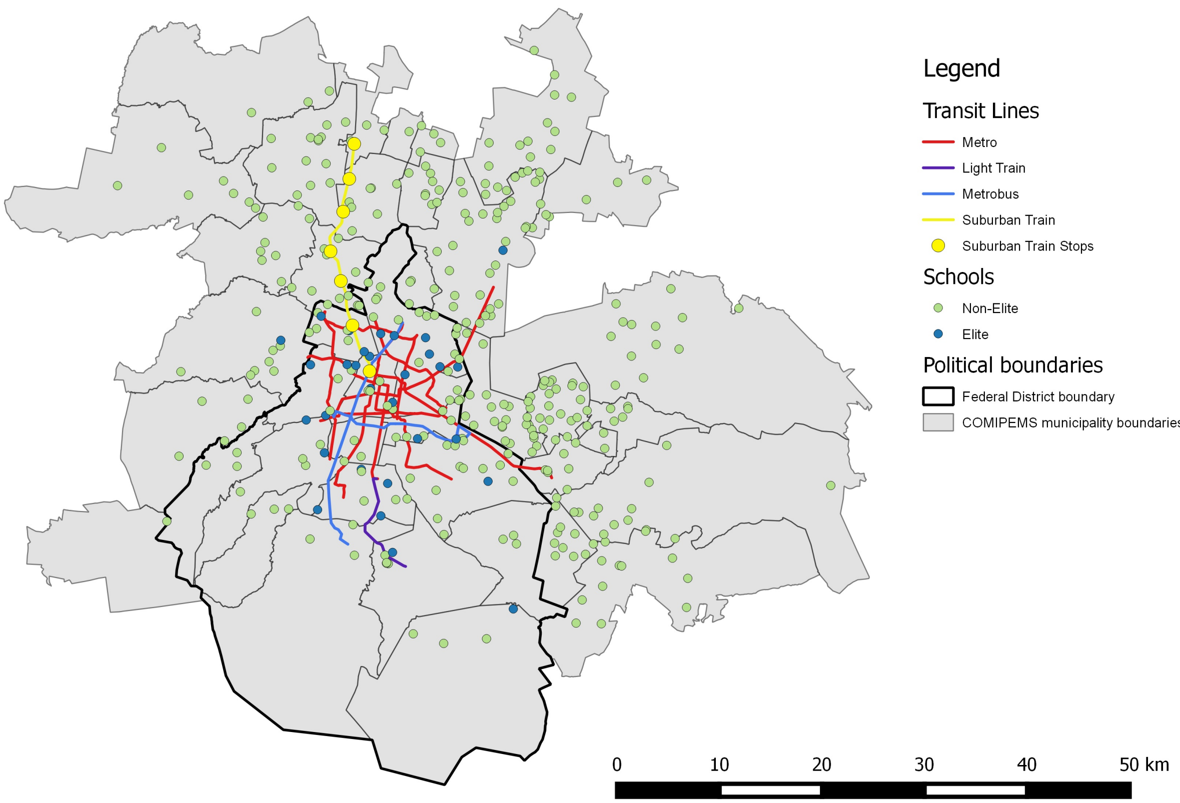
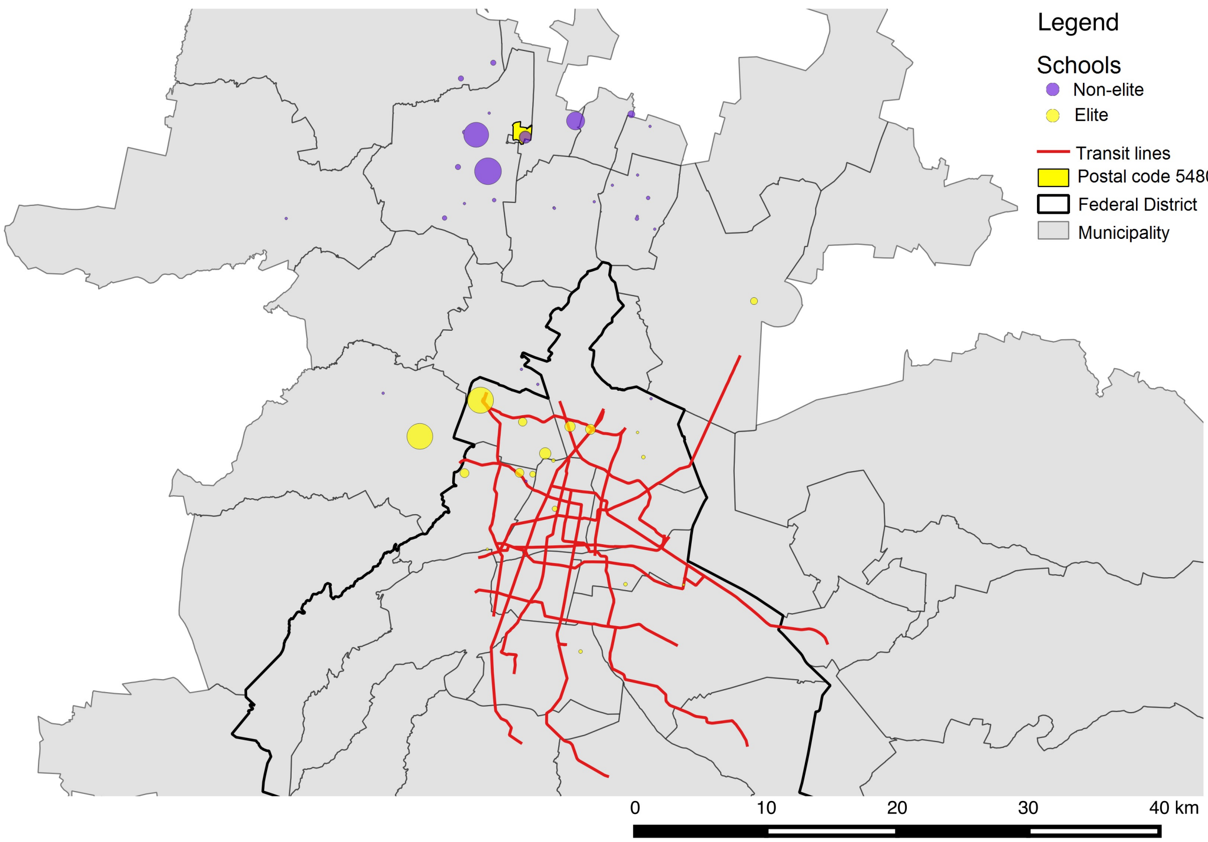


Figure 2: First choice schools for selected suburban postal code prior to Suburban Train (2007)



Note: Area of circles is proportional to the number of students from the selected postal code that listed it as their first choice.

Figure 3: Areas affected by Suburban Train: Treatment and Suburban Ring Control

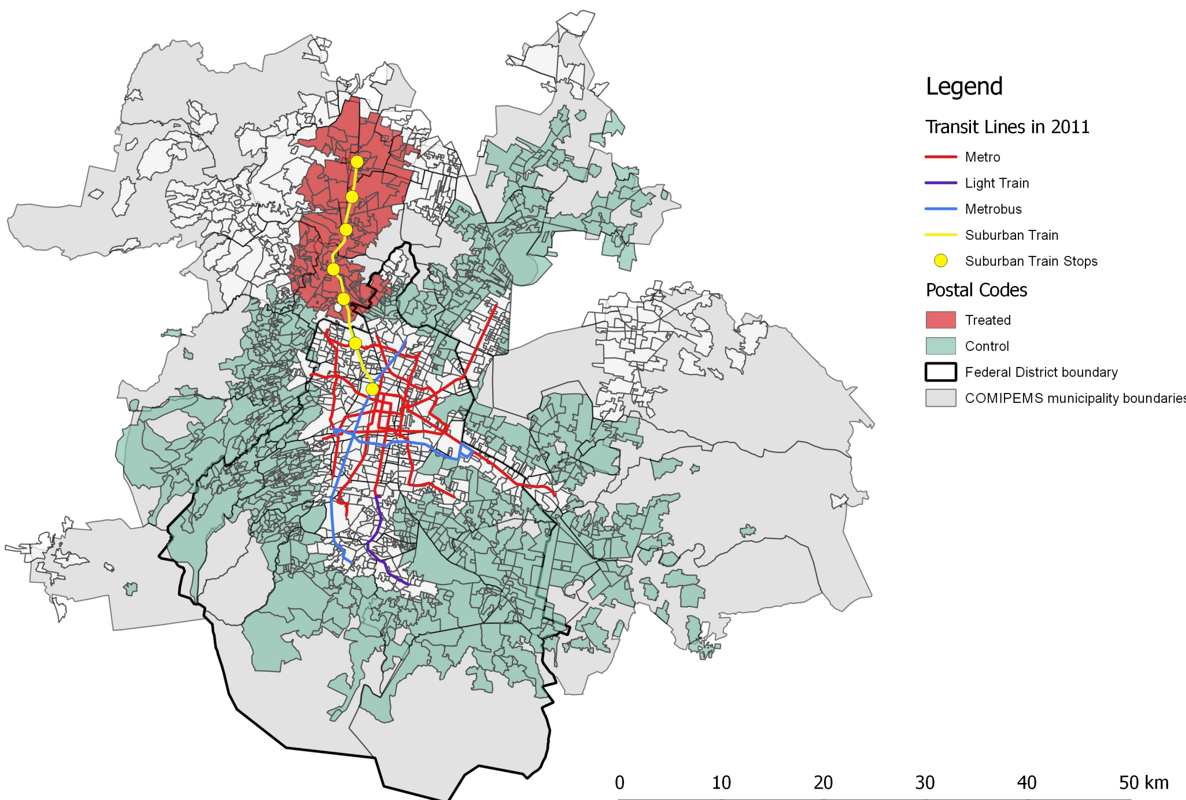


Figure 4: Areas affected by Suburban Train: Treatment and Planned Stations Control

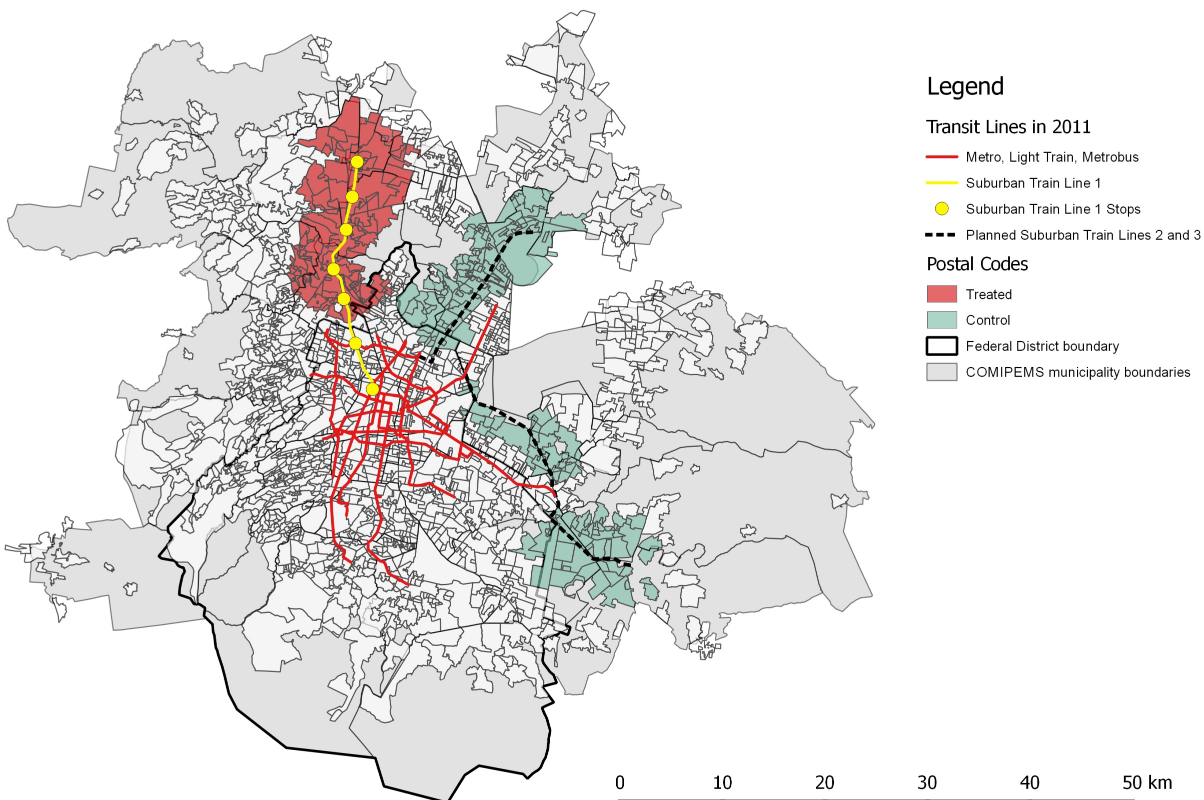
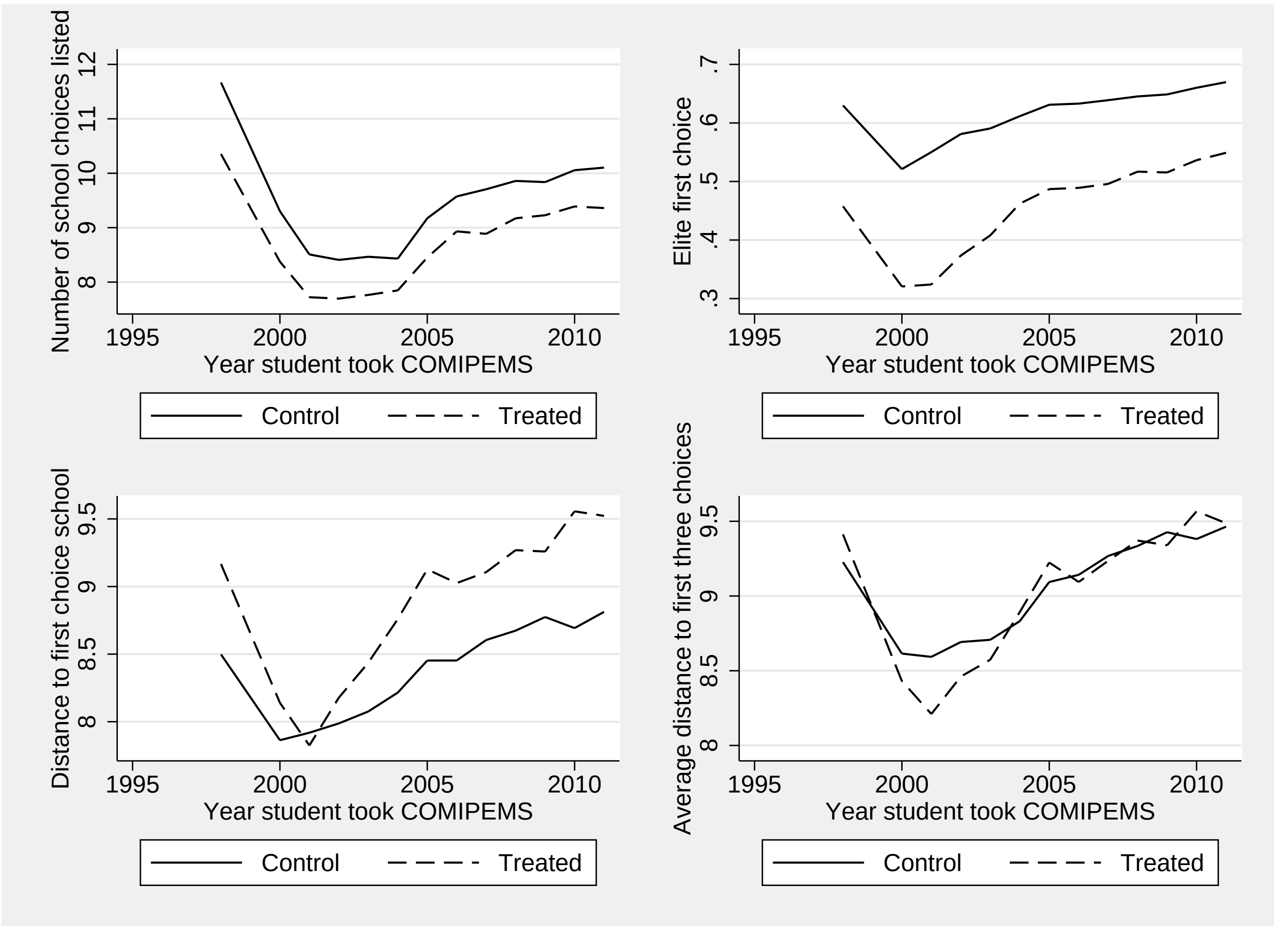
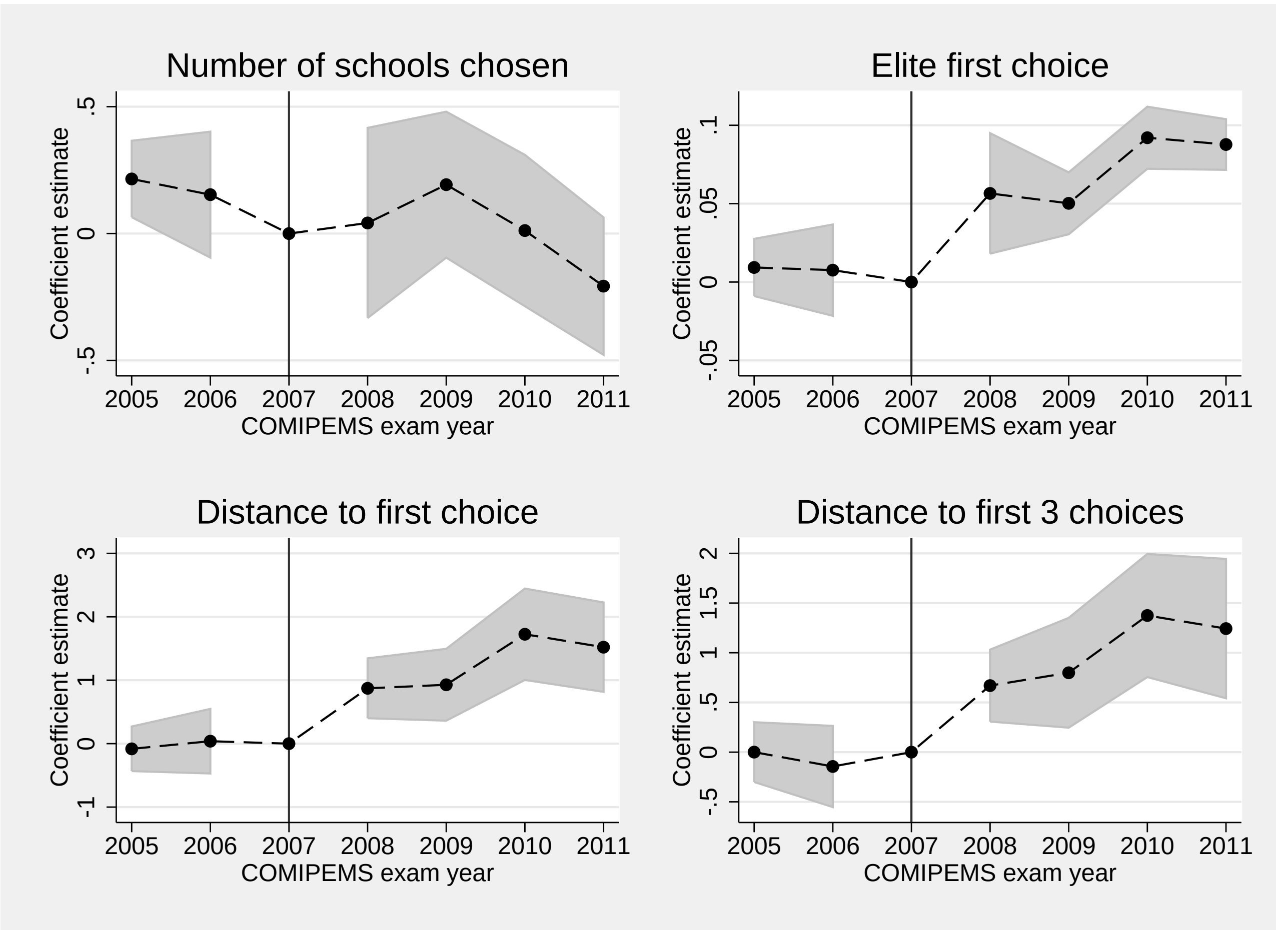


Figure 5: Trends in key outcomes in Suburban Train-affected and control areas, 1998-2011



Note: Treatment group is limited to students who are within 5 kilometers of a Suburban Train station, since this is the criterion used to define the treatment sample in most of the empirical analysis. The control sample here refers to the suburban ring control.

Figure 6: Event study of Suburban Train introduction for distant, high-scoring, high-SES students



Note: Sample is limited to high-scoring, high-SES students who are either in the suburban ring control group or the subset of the treatment group that experiences an above-median decrease in distance to transit station due to the Suburban Train introduction. Coefficients are year-specific treatment effects for the binary treatment indicator. Shaded areas are 95% confidence intervals. The final pre-treatment year is 2007.

Tables

Table 1: Summary statistics, students choosing schools in 2007 COMIPEMS exam year (pre-expansion)

	(1) Treatment	(2) Suburban ring control	(3) Planned station control
Outcomes			
Number of school choices listed	8.89	9.71	10.10
	[3.31]	[3.69]	[3.70]
Elite first choice	0.50	0.64	0.56
	[0.50]	[0.48]	[0.50]
Distance to first choice school	9.11	8.60	8.80
	[6.59]	[6.55]	[7.28]
Average distance to first three choices	9.24	9.27	9.34
	[5.73]	[5.77]	[6.54]
Distance to assigned school	7.83	7.49	7.32
	[6.27]	[6.42]	[6.90]
Cut-off score, assigned school	−0.45	−0.52	−0.57
	[0.87]	[0.96]	[0.92]
Elite assigned school	0.25	0.25	0.20
	[0.43]	[0.43]	[0.40]
Transit access			
Distance from closest transit station (km)	11.81	6.41	5.93
	[5.59]	[3.92]	[3.18]
Post-train change in distance to closest station (km)	−8.94	0.00	0.00
	[5.45]	[0.00]	[0.00]
Student characteristics			
Normalized COMIPEMS score	0.02	−0.05	−0.11
	[0.98]	[0.98]	[0.95]
Parental education (years)	10.67	10.07	9.90
	[3.33]	[3.31]	[3.25]
Male	0.48	0.48	0.48
	[0.50]	[0.50]	[0.50]
Oldest sibling	0.36	0.36	0.36
	[0.48]	[0.48]	[0.48]
Only child	0.05	0.05	0.05
	[0.22]	[0.22]	[0.21]
Number of siblings	2.04	2.21	2.26
	[1.33]	[1.44]	[1.46]
Private middle school	0.08	0.05	0.04
	[0.27]	[0.21]	[0.20]
Number of schools within 2.5km	2.45	3.44	4.86
	[1.58]	[2.47]	[2.25]
Observations	13838	93373	36794

Note: Sample is limited to students taking the COMIPEMS exam in 2007, the year before the Suburban Train began service. Standard deviations are in brackets.

Table 2: Effect of Suburban Train opening on student demographics

Panel A. Suburban ring control group	(1) Male	(2) Max parental education	(3) Private middle school	(4) Normalized COMIPEMS score	(5) Oldest sibling	(6) Only child	(7) Sibship size
Distance change x post	−0.0124	−0.0443	−0.0059	−0.0195	−0.0071	0.0017	−0.0039
	(0.0057)	(0.0563)	(0.0044)	(0.0174)	(0.0031)	(0.0030)	(0.0206)
	[0.12]	[0.52]	[0.26]	[0.32]	[0.13]	[0.61]	[0.69]
Observations	774606	774606	774606	774606	774606	774606	774606
Adjusted R^2	0.001	0.105	0.049	0.038	0.003	0.007	0.043
Dep. var. mean (2007)	0.480	10.147	0.051	−0.043	0.361	0.049	2.185

Panel B. Planned station control group	(1) Male	(2) Max parental education	(3) Private middle school	(4) Normalized COMIPEMS score	(5) Oldest sibling	(6) Only child	(7) Sibship size
Distance change x post	−0.012	−0.065	−0.006	−0.023	−0.007	0.001	−0.006
	(0.006)	(0.057)	(0.004)	(0.017)	(0.004)	(0.003)	(0.022)
	[0.13]	[0.35]	[0.23]	[0.31]	[0.23]	[0.66]	[0.68]
Observations	364380	364380	364380	364380	364380	364380	364380
Adjusted R^2	0.000	0.106	0.035	0.029	0.005	0.006	0.044
Dep. var. mean (2007)	0.481	10.114	0.052	−0.072	0.360	0.047	2.198

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, and municipality-by-treatment linear time trends.

Standard errors clustered at the municipality-by-treatment area level in parentheses. Wild bootstrapped p-values in brackets. * p < 0.1, ** p < 0.05, *** p < 0.01, according to wild p-values.

Table 3: Effect of Suburban Train opening on school choice outcomes

Panel A. Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices	(5) Number of schools chosen	(6) Elite first choice	(7) Distance to first choice	(8) Distance to first 3 choices
Close to station x post	0.135	0.004	0.109	0.158*				
	(0.113)	(0.007)	(0.066)	(0.081)				
	[0.31]	[0.56]	[0.16]	[0.06]				
Distance change x post					0.162	0.007	0.174*	0.186
					(0.123)	(0.006)	(0.060)	(0.074)
					[0.27]	[0.34]	[0.08]	[0.11]
Observations	774606	774606	774606	774606	774606	774606	774606	774606
Adjusted R^2	0.087	0.215	0.192	0.221	0.087	0.215	0.192	0.221
Dep. var. mean (2007)	9.602	0.621	8.672	9.267	9.602	0.621	8.672	9.267

Panel B. Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices	(5) Number of schools chosen	(6) Elite first choice	(7) Distance to first choice	(8) Distance to first 3 choices
Close to station x post	0.087	0.009	0.059	0.131				
	(0.128)	(0.008)	(0.075)	(0.084)				
	[0.50]	[0.37]	[0.45]	[0.18]				
Distance changed x post					0.134	0.010	0.149*	0.171*
					(0.134)	(0.006)	(0.066)	(0.076)
					[0.31]	[0.19]	[0.10]	[0.09]
Observations	364380	364380	364380	364380	364380	364380	364380	364380
Adjusted R^2	0.074	0.170	0.143	0.168	0.074	0.170	0.143	0.168
Dep. var. mean (2007)	9.772	0.541	8.883	9.313	9.772	0.541	8.883	9.313

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, and covariates (male, years of parental education, private middle school, normalized COMIPEMS score, number of schools within 2.5 kilometers, oldest sibling, only child, and number of siblings).

Standard errors clustered at the municipality-by-treatment area level in parentheses. Wild bootstrapped p-values in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, according to wild p-values.

Table 4: Effects on school choice outcomes, heterogeneity by parental education and COMIPEMS score

Panel A. Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.092	−0.009	−0.101	−0.021
	(0.127)	(0.010)	(0.102)	(0.113)
	[0.48]	[0.45]	[0.44]	[0.81]
Low ed/High score x distance change x post	0.105	−0.011	−0.121	−0.033
	(0.171)	(0.008)	(0.126)	(0.088)
	[0.60]	[0.36]	[0.57]	[0.77]
High ed/Low score x distance change x post	0.253**	0.001	0.097	0.134
	(0.100)	(0.012)	(0.218)	(0.133)
	[0.05]	[0.95]	[0.65]	[0.33]
High ed/High score x distance change x post	0.190	0.029***	0.557***	0.492***
	(0.151)	(0.007)	(0.090)	(0.073)
	[0.25]	[0.00]	[0.00]	[0.00]
Observations	774606	774606	774606	774606
Adjusted R^2	0.093	0.222	0.204	0.238
Dep. var. mean, low ed/low score (2007)	9.366	0.473	7.682	8.153
Dep. var. mean, low ed/high score (2007)	9.649	0.645	8.930	9.561
Dep. var. mean, high ed/low score (2007)	9.840	0.617	8.673	9.192
Dep. var. mean, high ed/high score (2007)	9.711	0.787	9.690	10.459

Panel B. Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.064	−0.004	−0.180	−0.085
	(0.136)	(0.010)	(0.110)	(0.114)
	[0.72]	[0.74]	[0.17]	[0.53]
Low ed/High score x distance change x post	0.119	−0.003	−0.056	0.014
	(0.182)	(0.009)	(0.126)	(0.097)
	[0.49]	[0.68]	[0.66]	[0.98]
High ed/Low score x distance change x post	0.207	−0.003	0.006	0.114
	(0.117)	(0.013)	(0.223)	(0.143)
	[0.19]	[0.77]	[0.88]	[0.49]
High ed/High score x distance change x post	0.132	0.030***	0.538***	0.464***
	(0.157)	(0.008)	(0.150)	(0.100)
	[0.54]	[0.00]	[0.00]	[0.00]
Observations	364380	364380	364380	364380
Adjusted R^2	0.079	0.176	0.153	0.182
Dep. var. mean, low ed/low score (2007)	9.534	0.400	7.567	7.920
Dep. var. mean, low ed/high score (2007)	9.757	0.571	9.069	9.520
Dep. var. mean, high ed/low score (2007)	10.034	0.536	8.977	9.294
Dep. var. mean, high ed/high score (2007)	9.922	0.704	10.378	10.960

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

Standard errors clustered at the municipality-by-treatment area level in parentheses. Wild bootstrapped p-values in brackets. * p < 0.1, ** p < 0.05, *** p < 0.01, according to wild p-values.

Table 5: Effects on commute-accessible outcomes

Suburban ring control group	(1) Number of commute schools	(2) First choice, commute	(3) First choice, distance to city center	(4) First 3 choices, distance to city center
Low ed/Low score x distance change x post	−0.022	−0.008	0.126	0.065
	(0.056)	(0.011)	(0.093)	(0.066)
	[0.64]	[0.49]	[0.23]	[0.41]
Low ed/High score x distance change x post	−0.125**	−0.020**	0.325	0.195
	(0.044)	(0.006)	(0.163)	(0.087)
	[0.03]	[0.02]	[0.14]	[0.12]
High ed/Low score x distance change x post	0.039	−0.017	0.112	−0.036
	(0.074)	(0.011)	(0.183)	(0.114)
	[0.62]	[0.32]	[0.61]	[0.73]
High ed/High score x distance change x post	0.081	0.018	−0.432**	−0.314***
	(0.069)	(0.011)	(0.120)	(0.065)
	[0.25]	[0.16]	[0.04]	[0.01]
Observations	774606	774606	774606	774606
Adjusted R^2	0.369	0.240	0.388	0.486
Dep. var. mean, low ed/low score (2007)	3.993	0.434	15.453	15.720
Dep. var. mean, low ed/high score (2007)	4.957	0.539	13.902	13.994
Dep. var. mean, high ed/low score (2007)	4.958	0.535	14.014	14.253
Dep. var. mean, high ed/high score (2007)	5.888	0.641	12.406	12.391

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

Standard errors clustered at the municipality-by-treatment area level in parentheses. Wild bootstrapped p-values in brackets. * p < 0.1, ** p < 0.05, *** p < 0.01, according to wild p-values.

Table 6: Effects on assigned school outcomes

Suburban ring control group	(1) Assigned school cutoff (2007)	(2) Elite assigned choice	(3) Distance to assigned school	(4) Assigned school commute accessible
Low ed/Low score x distance change x post	0.005		0.011	0.019
	(0.026)		(0.143)	(0.010)
	[0.88]		[0.97]	[0.18]
Low ed/High score x distance change x post	0.010	−0.008	−0.288	−0.002
	(0.023)	(0.012)	(0.208)	(0.014)
	[0.70]	[0.57]	[0.29]	[0.97]
High ed/Low score x distance change x post	−0.001		0.208	0.008
	(0.017)		(0.197)	(0.007)
	[0.97]		[0.35]	[0.41]
High ed/High score x distance change x post	0.058***	0.020**	0.238**	0.008
	(0.011)	(0.005)	(0.069)	(0.011)
	[0.00]	[0.03]	[0.02]	[0.67]
Observations	537677	339124	595513	595513
Adjusted R^2	0.598	0.479	0.168	0.175
Dep. var. mean, low ed/low score (2007)	−1.364		6.053	0.111
Dep. var. mean, low ed/high score (2007)	−0.186	0.325	8.002	0.234
Dep. var. mean, high ed/low score (2007)	−1.213		6.348	0.140
Dep. var. mean, high ed/high score (2007)	0.253	0.515	9.020	0.324

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

Standard errors clustered at the municipality-by-treatment area level in parentheses. Wild bootstrapped p-values in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, according to wild p-values.

Table 7: Instrumental variables estimates, straight line distance to historical train route

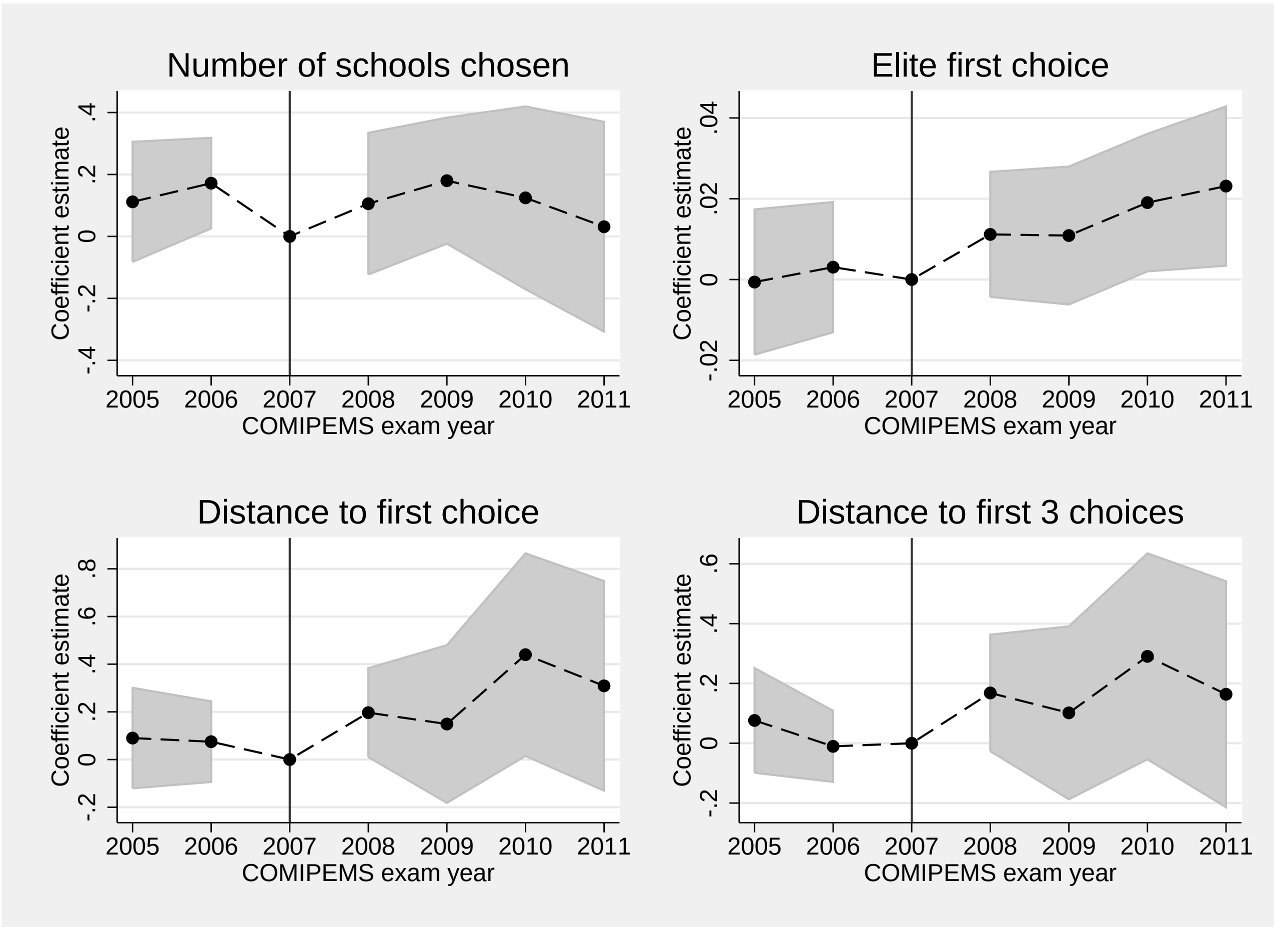
Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.088	−0.009	−0.106	−0.028
	(0.122)	(0.011)	(0.107)	(0.118)
Low ed/High score x distance change x post	0.095	−0.013	−0.139	−0.045
	(0.163)	(0.009)	(0.125)	(0.088)
High ed/Low score x distance change x post	0.251**	0.000	0.097	0.130
	(0.095)	(0.012)	(0.225)	(0.138)
High ed/High score x distance change x post	0.196	0.028***	0.537***	0.485***
	(0.149)	(0.007)	(0.093)	(0.075)
Cragg-Donald Wald F	37773788	37773788	37773788	37773788
Observations	774606	774606	774606	774606
Overall R^2	0.093	0.222	0.204	0.238
Dep. var. mean, low ed/low score (2007)	9.366	0.473	7.682	8.153
Dep. var. mean, low ed/high score (2007)	9.649	0.645	8.930	9.561
Dep. var. mean, high ed/low score (2007)	9.840	0.617	8.673	9.192
Dep. var. mean, high ed/high score (2007)	9.711	0.787	9.690	10.459

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Appendix (not for publication)

Figure A.1: Event study of Suburban Train introduction



Note: The suburban ring control group is used. Coefficients are year-specific treatment effects for the binary treatment indicator. Shaded areas are 95% confidence intervals. The final pre-treatment year is 2007. Event study regressions include postal code fixed effects, COMIPEMS exam year fixed effects, covariates, and treatment-by-year interaction terms, excluding 2007.

Table A.1: Summary statistics, students choosing schools in 2007 COMIPEMS exam year

	(1) Treatment	(2) Suburban ring control	(3) Planned station control	(4) City center	(5) Treatment - Suburban ring	(6) Treatment - Planned station
Outcomes						
Number of school choices listed	8.89	9.71	10.10	9.91	−0.82***	−1.21***
	[3.31]	[3.69]	[3.70]	[3.94]	(0.03)	(0.03)
Elite first choice	0.50	0.64	0.56	0.85	−0.14***	−0.06***
	[0.50]	[0.48]	[0.50]	[0.36]	(0.00)	(0.00)
Distance to first choice school	9.11	8.60	8.80	5.08	0.50***	0.31***
	[6.59]	[6.55]	[7.28]	[3.82]	(0.06)	(0.07)
Average distance to first three choices	9.24	9.27	9.34	6.00	−0.03	−0.10
	[5.73]	[5.77]	[6.54]	[3.04]	(0.05)	(0.06)
Distance to assigned school	7.83	7.49	7.32	5.61	0.35***	0.51***
	[6.27]	[6.42]	[6.90]	[4.43]	(0.07)	(0.07)
Cut-off score, assigned school	−0.45	−0.52	−0.57	−0.05	0.07***	0.13***
	[0.87]	[0.96]	[0.92]	[1.03]	(0.01)	(0.01)
Elite assigned school	0.25	0.25	0.20	0.44	0.00	0.04***
	[0.43]	[0.43]	[0.40]	[0.50]	(0.00)	(0.00)
Transit access						
Distance from closest transit station (km)	11.81	6.41	5.93	0.83	5.39***	5.87***
	[5.59]	[3.92]	[3.18]	[0.43]	(0.05)	(0.05)
Post-train change in distance to closest station (km)	−8.94	0.00	0.00	0.00	−8.94***	−8.94***
	[5.45]	[0.00]	[0.00]	[0.00]	(0.05)	(0.05)
Student characteristics						
Normalized COMIPEMS score	0.02	−0.05	−0.11	0.26	0.07***	0.12***
	[0.98]	[0.98]	[0.95]	[1.04]	(0.01)	(0.01)
Parental education (years)	10.67	10.07	9.90	11.68	0.60***	0.77***
	[3.33]	[3.31]	[3.25]	[3.36]	(0.03)	(0.03)
Male	0.48	0.48	0.48	0.48	0.00	0.00
	[0.50]	[0.50]	[0.50]	[0.50]	(0.00)	(0.00)
Oldest sibling	0.36	0.36	0.36	0.36	0.00	0.00
	[0.48]	[0.48]	[0.48]	[0.48]	(0.00)	(0.00)
Only child	0.05	0.05	0.05	0.09	0.00	0.01**
	[0.22]	[0.22]	[0.21]	[0.29]	(0.00)	(0.00)
Number of siblings	2.04	2.21	2.26	1.78	−0.16***	−0.21***
	[1.33]	[1.44]	[1.46]	[1.24]	(0.01)	(0.01)
Private middle school	0.08	0.05	0.04	0.13	0.03***	0.03***
	[0.27]	[0.21]	[0.20]	[0.33]	(0.00)	(0.00)
Number of schools within 2.5km	2.45	3.44	4.86	3.60	−0.99***	−2.41***
	[1.58]	[2.47]	[2.25]	[1.99]	(0.02)	(0.02)
Observations	13838	93373	36794	45160	107211	50632

Note: Sample is limited to students taking the COMIPEMS exam in 2007, the year before the Suburban Train began service. Standard deviations are in brackets. Standard errors for t-test of difference in means between treatment and control samples are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Comparison of AIC values for models explaining outcomes between 2001 and 2007, treatment and suburban ring control groups

	Treatment group				Control group			
	(1) Elite first choice	(2) Number of schools chosen	(3) Distance to first choice	(4) Distance to first 3 choices	(5) Elite first choice	(6) Number of schools chosen	(7) Distance to first choice	(8) Distance to first 3 choices
No trend break	458235.7	116373.7	561282.5	533592.3	3064351.4	702534.2	3647061.0	3469994.2
Trend break in 2003	458042.9	116308.2	561240.8	533570.4	3062334.5	702357.4	3647010.9	3469903.4
Trend break in 2004	458037.7	116271.0	561211.3	533535.1	3062036.2	702382.4	3646963.8	3469886.5
Trend break in 2005	**457921.9**	**116270.0**	561216.5	533545.4	**3061739.9**	702362.0	3646997.9	**3469872.4**
Trend break in 2006	458056.8	116270.4	**561203.0**	**533531.6**	3062964.5	**702325.9**	**3646960.4**	3469929.0

Note: Rows indicate model. Columns indicate outcome. All specifications include postal code fixed effects and municipality-by-treatment linear trends in COMIPEMS exam year. Rows indicating a trend break allow for an intercept shift in the stated year and for the municipality-by-treatment linear trends in COMIPEMS exam year to differ starting in that year. Bold-face values indicate the AIC-optimal model for that outcome.

Table A.3: Effects on nonelite school choice outcomes

Panel A. Suburban ring control group	(1) Number of nonelite choices	(2) 2007 cutoff for first nonelite	(3) First nonelite is commute accessible	(4) Distance to first nonelite choice
Low ed/Low score x distance change x post	0.090	0.028	0.010	0.023
	(0.147)	(0.020)	(0.007)	(0.098)
Low ed/High score x distance change x post	0.126	0.001	−0.008	−0.119
	(0.184)	(0.018)	(0.007)	(0.206)
High ed/Low score x distance change x post	0.205**	0.042**	−0.006	0.109
	(0.090)	(0.016)	(0.006)	(0.208)
High ed/High score x distance change x post	0.095	0.010	0.004	0.155
	(0.133)	(0.030)	(0.008)	(0.126)
Observations	774606	622984	705050	705050
Adjusted R^2	0.193	0.140	0.279	0.102
Dep. var. mean, low ed/low score (2007)	7.075	−0.678	0.295	6.158
Dep. var. mean, low ed/high score (2007)	6.056	−0.536	0.332	6.521
Dep. var. mean, high ed/low score (2007)	6.489	−0.465	0.341	6.411
Dep. var. mean, high ed/high score (2007)	4.768	−0.311	0.375	6.555

Panel B. Planned station control group	(1) Number of nonelite choices	(2) 2007 cutoff for first nonelite	(3) First nonelite is commute accessible	(4) Distance to first nonelite choice
Low ed/Low score x distance change x post	0.081	0.021	0.004	−0.059
	(0.152)	(0.021)	(0.008)	(0.103)
Low ed/High score x distance change x post	0.107	0.007	−0.002	−0.134
	(0.196)	(0.021)	(0.007)	(0.217)
High ed/Low score x distance change x post	0.209**	0.033	−0.010	0.091
	(0.100)	(0.022)	(0.006)	(0.213)
High ed/High score x distance change x post	0.098	0.005	0.004	0.071
	(0.146)	(0.033)	(0.010)	(0.127)
Observations	364380	293791	336555	336555
Adjusted R^2	0.142	0.146	0.153	0.061
Dep. var. mean, low ed/low score (2007)	7.608	−0.730	0.245	5.710
Dep. var. mean, low ed/high score (2007)	6.585	−0.576	0.266	6.089
Dep. var. mean, high ed/low score (2007)	7.133	−0.490	0.268	6.087
Dep. var. mean, high ed/high score (2007)	5.485	−0.330	0.273	6.210

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.4: Effects on commute-accessible outcomes, planned station control group

Planned station control group	(1) Number of commute schools	(2) First choice, commute	(3) First choice, distance to city center	(4) First 3 choices, distance to city center
Low ed/Low score x distance change x post	−0.053	−0.010	0.107	0.057
	(0.057)	(0.011)	(0.095)	(0.067)
	[0.31]	[0.35]	[0.26]	[0.44]
Low ed/High score x distance change x post	−0.102*	−0.014*	0.182	0.078
	(0.047)	(0.007)	(0.181)	(0.105)
	[0.06]	[0.10]	[0.40]	[0.50]
High ed/Low score x distance change x post	−0.004	−0.028*	0.268	0.028
	(0.076)	(0.012)	(0.199)	(0.114)
	[1.00]	[0.08]	[0.31]	[0.83]
High ed/High score x distance change x post	0.045	0.012	−0.444**	−0.346***
	(0.070)	(0.013)	(0.148)	(0.081)
	[0.60]	[0.37]	[0.05]	[0.01]
Observations	364380	364380	364380	364380
Adjusted R^2	0.325	0.173	0.344	0.434
Dep. var. mean, low ed/low score (2007)	3.613	0.431	16.462	16.871
Dep. var. mean, low ed/high score (2007)	4.512	0.530	14.712	14.951
Dep. var. mean, high ed/low score (2007)	4.478	0.526	14.930	15.303
Dep. var. mean, high ed/high score (2007)	5.454	0.614	13.101	13.187

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

Standard errors clustered at the municipality-by-treatment area level in parentheses. Wild bootstrapped p-values in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, according to wild p-values.

Table A.5: Average effects on assigned school choice outcomes

Panel A. Suburban ring control group	(1) Assigned school cutoff (2007)	(2) Elite assigned choice	(3) Distance to assigned school	(4) Assigned school commute accessible
Distance changed x post	0.030***	0.006*	0.061	0.008
	(0.007)	(0.003)	(0.039)	(0.006)
Observations	537677	595513	595513	595513
Adjusted R^2	0.591	0.501	0.140	0.160
Dep. var. mean (2007)	−0.530	0.246	7.530	0.214

Panel B. Planned station control group	(1) Assigned school cutoff (2007)	(2) Elite assigned choice	(3) Distance to assigned school	(4) Assigned school commute accessible
Distance changed x post	0.027***	0.007**	0.076	0.012*
	(0.008)	(0.003)	(0.050)	(0.006)
Observations	258554	289171	289171	289171
Adjusted R^2	0.540	0.448	0.143	0.172
Dep. var. mean (2007)	−0.606	0.216	7.456	0.212

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.6: Effects on assigned school outcomes, planned station control group

Planned station control group	(1) Assigned school cutoff (2007)	(2) Elite assigned choice	(3) Distance to assigned school	(4) Assigned school commute accessible
Low ed/Low score x distance change x post	−0.007		0.027	0.022**
	(0.027)		(0.138)	(0.010)
	[0.85]		[0.93]	[0.04]
Low ed/High score x distance change x post	0.030	−0.004	−0.321	0.003
	(0.022)	(0.012)	(0.216)	(0.014)
	[0.28]	[0.74]	[0.28]	[0.78]
High ed/Low score x distance change x post	−0.025		0.321	0.015
	(0.016)		(0.203)	(0.009)
	[0.25]		[0.13]	[0.10]
High ed/High score x distance change x post	0.063***	0.021***	0.229*	0.007
	(0.014)	(0.004)	(0.093)	(0.011)
	[0.00]	[0.01]	[0.08]	[0.71]
Observations	258554	158234	289171	289171
Adjusted R^2	0.547	0.419	0.165	0.187
Dep. var. mean, low ed/low score (2007)	−1.365		5.724	0.099
Dep. var. mean, low ed/high score (2007)	−0.267	0.294	7.907	0.248
Dep. var. mean, high ed/low score (2007)	−1.206		6.107	0.126
Dep. var. mean, high ed/high score (2007)	0.147	0.466	9.429	0.335

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

Standard errors clustered at the municipality-by-treatment area level in parentheses. Wild bootstrapped p-values in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, according to wild p-values.

Table A.7: Instrumental variables first stage

Panel A. Suburban ring control group	(1) Low ed/Low score x distance change x post	(2) Low ed/High score x distance change x post	(3) High ed/Low score x distance change x post	(4) High ed/High score x distance change x post
Low ed/Low score x distance change to line x post	0.983***	−0.000	−0.000	−0.000
	(0.010)	(0.000)	(0.000)	(0.000)
Low ed/High score x distance change to line x post	−0.000	0.984***	0.000	0.000
	(0.000)	(0.009)	(0.000)	(0.000)
High ed/Low score x distance change to line x post	−0.000	0.000	0.988***	0.000
	(0.000)	(0.000)	(0.008)	(0.000)
High ed/High score x distance change to line x post	0.000	0.000	−0.000	0.990***
	(0.000)	(0.000)	(0.000)	(0.008)
Observations	774606	774606	774606	774606
Adjusted R^2	0.999	0.999	1.000	1.000

Panel B. Planned station control group	(1) Low ed/Low score x distance change x post	(2) Low ed/High score x distance change x post	(3) High ed/Low score x distance change x post	(4) High ed/High score x distance change x post
Low ed/Low score x distance change to line x post	0.985***	0.000	0.000	−0.000
	(0.010)	(0.000)	(0.000)	(0.000)
Low ed/High score x distance change to line x post	0.000	0.985***	−0.000	0.000
	(0.000)	(0.009)	(0.000)	(0.000)
High ed/Low score x distance change to line x post	0.000	−0.000	0.990***	0.000
	(0.000)	(0.000)	(0.008)	(0.000)
High ed/High score x distance change to line x post	−0.000	0.000	0.000	0.992***
	(0.000)	(0.000)	(0.000)	(0.008)
Observations	364380	364380	364380	364380
Adjusted R^2	0.999	0.999	1.000	1.000

Note: Distance change to line measures the distance between the postal code centroid and the Suburban Train line (in contrast to the station). First stage results are identical for all school choice outcomes. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.8: Instrumental variables estimates, straight line distance to Suburban Train line, planned station control group

Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.057	−0.004	−0.187	−0.093
	(0.131)	(0.011)	(0.116)	(0.119)
Low ed/High score x distance change x post	0.108	−0.005	−0.075	0.001
	(0.173)	(0.009)	(0.126)	(0.097)
High ed/Low score x distance change x post	0.203*	−0.003	0.004	0.109
	(0.111)	(0.014)	(0.231)	(0.148)
High ed/High score x distance change x post	0.136	0.029***	0.515***	0.454***
	(0.156)	(0.008)	(0.153)	(0.102)
Cragg-Donald Wald F	16260318	16260318	16260318	16260318
Observations	364380	364380	364380	364380
Overall R^2	0.079	0.176	0.153	0.182
Dep. var. mean, low ed/low score (2007)	9.534	0.400	7.567	7.920
Dep. var. mean, low ed/high score (2007)	9.757	0.571	9.069	9.520
Dep. var. mean, high ed/low score (2007)	10.034	0.536	8.977	9.294
Dep. var. mean, high ed/high score (2007)	9.922	0.704	10.378	10.960

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.9: Effects on school choice outcomes, three kilometer treatment buffer

Panel A. Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.156	−0.005	0.047	0.078
	(0.128)	(0.012)	(0.086)	(0.109)
Low ed/High score x distance change x post	0.045	−0.006	0.043	0.073
	(0.198)	(0.009)	(0.085)	(0.083)
High ed/Low score x distance change x post	0.168***	0.009	0.218	0.160
	(0.061)	(0.012)	(0.272)	(0.189)
High ed/High score x distance change x post	0.136	0.037***	0.683***	0.617***
	(0.146)	(0.006)	(0.113)	(0.115)
Observations	734722	734722	734722	734722
Adjusted R^2	0.093	0.220	0.209	0.245
Dep. var. mean, low ed/low score (2007)	9.379	0.479	7.675	8.161
Dep. var. mean, low ed/high score (2007)	9.671	0.652	8.909	9.564
Dep. var. mean, high ed/low score (2007)	9.881	0.628	8.658	9.213
Dep. var. mean, high ed/high score (2007)	9.770	0.804	9.650	10.470

Panel B. Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.206	0.004	0.012	0.069
	(0.129)	(0.014)	(0.089)	(0.116)
Low ed/High score x distance change x post	0.062	−0.009	0.035	0.102
	(0.211)	(0.010)	(0.087)	(0.089)
High ed/Low score x distance change x post	0.103	0.005	0.175	0.270
	(0.089)	(0.015)	(0.293)	(0.226)
High ed/High score x distance change x post	0.125	0.033***	0.557***	0.492***
	(0.158)	(0.006)	(0.120)	(0.115)
Observations	242130	242130	242130	242130
Adjusted R^2	0.081	0.167	0.149	0.187
Dep. var. mean, low ed/low score (2007)	9.627	0.436	7.268	7.722
Dep. var. mean, low ed/high score (2007)	9.830	0.604	8.516	9.081
Dep. var. mean, high ed/low score (2007)	10.148	0.582	8.299	8.790
Dep. var. mean, high ed/high score (2007)	10.090	0.757	9.526	10.230

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students in the main specification. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.10: Effects on school choice outcomes, seven kilometer treatment buffer

Panel A. Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.206	−0.009	−0.086	−0.016
	(0.150)	(0.009)	(0.075)	(0.104)
Low ed/High score x distance change x post	0.268	0.005	0.141	0.184
	(0.220)	(0.009)	(0.142)	(0.138)
High ed/Low score x distance change x post	0.399**	0.002	0.140	0.101
	(0.170)	(0.012)	(0.240)	(0.192)
High ed/High score x distance change x post	0.390**	0.022***	0.438***	0.434***
	(0.178)	(0.007)	(0.082)	(0.072)
Observations	818169	818169	818169	818169
Adjusted R^2	0.093	0.222	0.198	0.231
Dep. var. mean, low ed/low score (2007)	9.352	0.468	7.692	8.136
Dep. var. mean, low ed/high score (2007)	9.627	0.637	8.948	9.545
Dep. var. mean, high ed/low score (2007)	9.799	0.605	8.686	9.165
Dep. var. mean, high ed/high score (2007)	9.680	0.772	9.745	10.464

Panel B. Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.150	−0.007	−0.149*	−0.076
	(0.161)	(0.010)	(0.086)	(0.108)
Low ed/High score x distance change x post	0.257	0.009	0.180	0.210
	(0.232)	(0.010)	(0.138)	(0.144)
High ed/Low score x distance change x post	0.343*	−0.001	0.072	0.069
	(0.185)	(0.013)	(0.247)	(0.200)
High ed/High score x distance change x post	0.332*	0.024***	0.469***	0.457***
	(0.183)	(0.008)	(0.139)	(0.095)
Observations	460617	460617	460617	460617
Adjusted R^2	0.081	0.176	0.147	0.178
Dep. var. mean, low ed/low score (2007)	9.440	0.380	7.662	7.954
Dep. var. mean, low ed/high score (2007)	9.688	0.551	9.251	9.639
Dep. var. mean, high ed/low score (2007)	9.917	0.514	9.140	9.389
Dep. var. mean, high ed/high score (2007)	9.857	0.680	10.620	11.128

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students in the main specification. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.11: Effects on school choice outcomes, including Texcoco municipality

	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.102	−0.009	−0.105	−0.024
	(0.126)	(0.010)	(0.102)	(0.112)
Low ed/High score x distance change x post	0.119	−0.012	−0.146	−0.052
	(0.170)	(0.008)	(0.129)	(0.090)
High ed/Low score x distance change x post	0.263**	−0.000	0.060	0.109
	(0.100)	(0.012)	(0.219)	(0.134)
High ed/High score x distance change x post	0.191	0.028***	0.502***	0.449***
	(0.150)	(0.007)	(0.102)	(0.082)
Observations	791308	791308	791308	791308
Adjusted R^2	0.111	0.236	0.201	0.236
Dep. var. mean, low ed/low score (2007)	9.290	0.464	7.706	8.172
Dep. var. mean, low ed/high score (2007)	9.575	0.634	8.945	9.570
Dep. var. mean, high ed/low score (2007)	9.751	0.606	8.671	9.198
Dep. var. mean, high ed/high score (2007)	9.604	0.773	9.695	10.468

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students in the main specification. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.12: Effects on school choice outcomes, excluding Federal District postal codes

Panel A. Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.052	−0.011	−0.156	−0.066
	(0.135)	(0.010)	(0.108)	(0.114)
Low ed/High score x distance change x post	0.111	−0.013	−0.153	−0.078
	(0.180)	(0.009)	(0.135)	(0.105)
High ed/Low score x distance change x post	0.235*	0.002	0.047	0.114
	(0.119)	(0.013)	(0.217)	(0.135)
High ed/High score x distance change x post	0.192	0.029***	0.546***	0.485***
	(0.159)	(0.008)	(0.119)	(0.087)
Observations	504549	504549	504549	504549
Adjusted R^2	0.093	0.173	0.168	0.193
Dep. var. mean, low ed/low score (2007)	9.170	0.394	7.740	8.120
Dep. var. mean, low ed/high score (2007)	9.484	0.551	9.364	9.833
Dep. var. mean, high ed/low score (2007)	9.743	0.519	9.202	9.581
Dep. var. mean, high ed/high score (2007)	9.726	0.684	10.920	11.552

Panel B. Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.064	−0.006	−0.184	−0.087
	(0.139)	(0.010)	(0.112)	(0.114)
Low ed/High score x distance change x post	0.137	−0.003	−0.065	0.006
	(0.182)	(0.009)	(0.128)	(0.100)
High ed/Low score x distance change x post	0.213*	−0.004	−0.015	0.103
	(0.121)	(0.014)	(0.223)	(0.145)
High ed/High score x distance change x post	0.139	0.032***	0.543***	0.465***
	(0.161)	(0.009)	(0.167)	(0.108)
Observations	339031	339031	339031	339031
Adjusted R^2	0.079	0.157	0.138	0.168
Dep. var. mean, low ed/low score (2007)	9.503	0.384	7.676	8.011
Dep. var. mean, low ed/high score (2007)	9.721	0.551	9.270	9.702
Dep. var. mean, high ed/low score (2007)	10.020	0.514	9.231	9.520
Dep. var. mean, high ed/high score (2007)	9.887	0.681	10.840	11.388

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students in the main specification. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.13: Effects on school choice outcomes, postal code level trends

Panel A. Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.100	−0.001	−0.001	0.056
	(0.124)	(0.008)	(0.082)	(0.121)
Low ed/High score x distance change x post	0.152	−0.007	−0.169	−0.008
	(0.126)	(0.008)	(0.142)	(0.113)
High ed/Low score x distance change x post	0.275***	0.002	0.088	0.148
	(0.084)	(0.011)	(0.194)	(0.100)
High ed/High score x distance change x post	0.201	0.027***	0.430***	0.386***
	(0.149)	(0.007)	(0.142)	(0.099)
Observations	774606	774606	774606	774606
Adjusted R^2	0.028	0.028	0.014	0.021
Dep. var. mean, low ed/low score (2007)	9.366	0.473	7.682	8.153
Dep. var. mean, low ed/high score (2007)	9.649	0.645	8.930	9.561
Dep. var. mean, high ed/low score (2007)	9.840	0.617	8.673	9.192
Dep. var. mean, high ed/high score (2007)	9.711	0.787	9.690	10.459

Panel B. Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.065	0.005	−0.097	−0.025
	(0.137)	(0.009)	(0.085)	(0.121)
Low ed/High score x distance change x post	0.176	0.002	−0.110	0.039
	(0.134)	(0.008)	(0.140)	(0.125)
High ed/Low score x distance change x post	0.220**	−0.002	−0.029	0.111
	(0.100)	(0.013)	(0.196)	(0.112)
High ed/High score x distance change x post	0.130	0.028***	0.382*	0.334**
	(0.153)	(0.008)	(0.187)	(0.123)
Observations	364380	364380	364380	364380
Adjusted R^2	0.024	0.030	0.017	0.025
Dep. var. mean, low ed/low score (2007)	9.534	0.400	7.567	7.920
Dep. var. mean, low ed/high score (2007)	9.757	0.571	9.069	9.520
Dep. var. mean, high ed/low score (2007)	10.034	0.536	8.977	9.294
Dep. var. mean, high ed/high score (2007)	9.922	0.704	10.378	10.960

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, postal code linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.14: Effects on school choice outcomes, no linear trends

Panel A. Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.006	−0.011	−0.291***	−0.281***
	(0.120)	(0.007)	(0.088)	(0.062)
Low ed/High score x distance change x post	0.065	0.002	0.221	0.090
	(0.132)	(0.009)	(0.151)	(0.131)
High ed/Low score x distance change x post	−0.029	−0.007	−0.001	−0.105
	(0.098)	(0.005)	(0.102)	(0.089)
High ed/High score x distance change x post	−0.070	0.049***	0.883***	0.762***
	(0.090)	(0.006)	(0.161)	(0.131)
Observations	774606	774606	774606	774606
Adjusted R^2	0.092	0.222	0.203	0.237
Dep. var. mean, low ed/low score (2007)	9.366	0.473	7.682	8.153
Dep. var. mean, low ed/high score (2007)	9.649	0.645	8.930	9.561
Dep. var. mean, high ed/low score (2007)	9.840	0.617	8.673	9.192
Dep. var. mean, high ed/high score (2007)	9.711	0.787	9.690	10.459

Panel B. Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.085	−0.013	−0.347***	−0.321***
	(0.122)	(0.008)	(0.093)	(0.072)
Low ed/High score x distance change x post	0.129	−0.002	0.062	−0.073
	(0.135)	(0.010)	(0.187)	(0.152)
High ed/Low score x distance change x post	0.038	−0.010**	−0.025	−0.121
	(0.102)	(0.004)	(0.110)	(0.099)
High ed/High score x distance change x post	−0.030	0.038***	0.689***	0.534***
	(0.091)	(0.006)	(0.184)	(0.161)
Observations	364380	364380	364380	364380
Adjusted R^2	0.078	0.175	0.152	0.181
Dep. var. mean, low ed/low score (2007)	9.534	0.400	7.567	7.920
Dep. var. mean, low ed/high score (2007)	9.757	0.571	9.069	9.520
Dep. var. mean, high ed/low score (2007)	10.034	0.536	8.977	9.294
Dep. var. mean, high ed/high score (2007)	9.922	0.704	10.378	10.960

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.15: Effects on school choice outcomes, wild cluster bootstrapped p-values

Panel A. Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.092	−0.009	−0.101	−0.021
	(0.127)	(0.010)	(0.102)	(0.113)
	[0.45]	[0.38]	[0.42]	[0.78]
Low ed/High score x distance change x post	0.105	−0.011	−0.121	−0.033
	(0.171)	(0.008)	(0.126)	(0.088)
	[0.69]	[0.29]	[0.42]	[0.73]
High ed/Low score x distance change x post	0.253*	0.001	0.097	0.134
	(0.100)	(0.012)	(0.218)	(0.133)
	[0.10]	[0.96]	[0.88]	[0.52]
High ed/High score x distance change x post	0.190	0.029**	0.557**	0.492***
	(0.151)	(0.007)	(0.090)	(0.073)
	[0.29]	[0.02]	[0.03]	[0.01]
Observations	774606	774606	774606	774606
Adjusted R^2	0.093	0.222	0.204	0.238
Dep. var. mean, low ed/low score (2007)	9.366	0.473	7.682	8.153
Dep. var. mean, low ed/high score (2007)	9.649	0.645	8.930	9.561
Dep. var. mean, high ed/low score (2007)	9.840	0.617	8.673	9.192
Dep. var. mean, high ed/high score (2007)	9.711	0.787	9.690	10.459

Panel B. Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.064	−0.004	−0.180*	−0.085
	(0.136)	(0.010)	(0.110)	(0.114)
	[0.61]	[0.71]	[0.10]	[0.57]
Low ed/High score x distance change x post	0.119	−0.003	−0.056	0.014
	(0.182)	(0.009)	(0.126)	(0.097)
	[0.64]	[0.74]	[0.64]	[0.84]
High ed/Low score x distance change x post	0.207	−0.003	0.006	0.114
	(0.117)	(0.013)	(0.223)	(0.143)
	[0.24]	[0.99]	[0.92]	[0.53]
High ed/High score x distance change x post	0.132	0.030**	0.538*	0.464**
	(0.157)	(0.008)	(0.150)	(0.100)
	[0.46]	[0.02]	[0.06]	[0.03]
Observations	364380	364380	364380	364380
Adjusted R^2	0.079	0.176	0.153	0.182
Dep. var. mean, low ed/low score (2007)	9.534	0.400	7.567	7.920
Dep. var. mean, low ed/high score (2007)	9.757	0.571	9.069	9.520
Dep. var. mean, high ed/low score (2007)	10.034	0.536	8.977	9.294
Dep. var. mean, high ed/high score (2007)	9.922	0.704	10.378	10.960

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy.

Standard errors clustered at the municipality-by-treatment area level in parentheses. Wild cluster bootstrapped p-values in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, according to wild cluster p-values.

Table A.16: Effects on school choice outcomes, 2007 placebo treatment

Panel A. Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	−0.373*	−0.013	−0.097	0.017
	(0.152)	(0.014)	(0.179)	(0.190)
	[0.09]	[0.43]	[0.67]	[0.98]
Low ed/High score x distance change x post	−0.355	−0.006	−0.042	0.112
	(0.215)	(0.019)	(0.223)	(0.182)
	[0.21]	[0.70]	[0.88]	[0.68]
High ed/Low score x distance change x post	−0.200	−0.033	−0.302	−0.133
	(0.102)	(0.028)	(0.342)	(0.286)
	[0.19]	[0.25]	[0.54]	[0.64]
High ed/High score x distance change x post	0.029	0.008	0.032	0.272
	(0.201)	(0.012)	(0.202)	(0.186)
	[0.94]	[0.53]	[0.86]	[0.30]
Observations	312971	312971	312971	312971
Adjusted R^2	0.096	0.205	0.193	0.228
Dep. var. mean, low ed/low score (2006)	9.325	0.475	7.642	8.127
Dep. var. mean, low ed/high score (2006)	9.570	0.641	8.661	9.396
Dep. var. mean, high ed/low score (2006)	9.664	0.614	8.556	9.066
Dep. var. mean, high ed/high score (2006)	9.527	0.775	9.582	10.299

Panel B. Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	−0.441*	−0.014	−0.206	−0.119
	(0.162)	(0.014)	(0.197)	(0.215)
	[0.07]	[0.61]	[0.64]	[0.93]
Low ed/High score x distance change x post	−0.376	−0.012	−0.127	−0.089
	(0.229)	(0.020)	(0.237)	(0.196)
	[0.20]	[0.66]	[0.65]	[0.89]
High ed/Low score x distance change x post	−0.328	−0.037	−0.543	−0.227
	(0.123)	(0.028)	(0.378)	(0.318)
	[0.23]	[0.29]	[0.82]	[0.60]
High ed/High score x distance change x post	−0.089	0.005	−0.136	0.099
	(0.216)	(0.013)	(0.227)	(0.239)
	[0.85]	[0.61]	[0.33]	[0.39]
Observations	147978	147978	147978	147978
Adjusted R^2	0.080	0.156	0.136	0.163
Dep. var. mean, low ed/low score (2006)	9.449	0.388	7.556	7.858
Dep. var. mean, low ed/high score (2006)	9.689	0.557	8.778	9.285
Dep. var. mean, high ed/low score (2006)	9.797	0.521	8.868	9.198
Dep. var. mean, high ed/high score (2006)	9.707	0.678	10.200	10.727

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students in the main specification. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.17: Effects on school choice outcomes, 2004 pre-period start

Panel A. Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.119	−0.008	−0.072	−0.005
	(0.126)	(0.010)	(0.095)	(0.095)
Low ed/High score x distance change x post	0.180	−0.012	−0.164	−0.053
	(0.160)	(0.010)	(0.113)	(0.085)
High ed/Low score x distance change x post	0.198**	−0.004	0.096	0.117
	(0.090)	(0.011)	(0.217)	(0.141)
High ed/High score x distance change x post	0.181	0.033***	0.602***	0.550***
	(0.155)	(0.006)	(0.071)	(0.062)
Observations	864283	864283	864091	864283
Adjusted R^2	0.103	0.220	0.202	0.236
Dep. var. mean, low ed/low score (2007)	9.364	0.473	7.677	8.147
Dep. var. mean, low ed/high score (2007)	9.648	0.646	8.928	9.558
Dep. var. mean, high ed/low score (2007)	9.836	0.617	8.669	9.189
Dep. var. mean, high ed/high score (2007)	9.710	0.788	9.688	10.455

Panel B. Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.107	−0.000	−0.137	−0.048
	(0.134)	(0.010)	(0.102)	(0.097)
Low ed/High score x distance change x post	0.208	−0.005	−0.067	0.021
	(0.169)	(0.010)	(0.115)	(0.103)
High ed/Low score x distance change x post	0.169	−0.005	0.004	0.090
	(0.102)	(0.012)	(0.220)	(0.152)
High ed/High score x distance change x post	0.124	0.036***	0.616***	0.559***
	(0.161)	(0.008)	(0.130)	(0.098)
Observations	407169	407169	407031	407169
Adjusted R^2	0.088	0.173	0.150	0.178
Dep. var. mean, low ed/low score (2007)	9.533	0.400	7.566	7.918
Dep. var. mean, low ed/high score (2007)	9.756	0.572	9.063	9.516
Dep. var. mean, high ed/low score (2007)	10.031	0.536	8.982	9.298
Dep. var. mean, high ed/high score (2007)	9.920	0.704	10.380	10.960

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students in the main specification. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.

Table A.18: Effects on school choice outcomes, 2006 pre-period start

Panel A. Suburban ring control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.125	−0.007	−0.101	0.001
	(0.135)	(0.010)	(0.111)	(0.129)
Low ed/High score x distance change x post	0.158	−0.010	−0.106	−0.018
	(0.147)	(0.008)	(0.128)	(0.090)
High ed/Low score x distance change x post	0.279**	0.002	0.082	0.112
	(0.106)	(0.013)	(0.228)	(0.128)
High ed/High score x distance change x post	0.184	0.024**	0.469***	0.413***
	(0.145)	(0.009)	(0.126)	(0.097)
Observations	673840	673840	673840	673840
Adjusted R^2	0.090	0.228	0.207	0.241
Dep. var. mean, low ed/low score (2007)	9.366	0.473	7.684	8.154
Dep. var. mean, low ed/high score (2007)	9.648	0.645	8.931	9.561
Dep. var. mean, high ed/low score (2007)	9.840	0.616	8.673	9.191
Dep. var. mean, high ed/high score (2007)	9.710	0.787	9.697	10.465

Panel B. Planned station control group	(1) Number of schools chosen	(2) Elite first choice	(3) Distance to first choice	(4) Distance to first 3 choices
Low ed/Low score x distance change x post	0.086	−0.002	−0.170	−0.058
	(0.146)	(0.011)	(0.119)	(0.131)
Low ed/High score x distance change x post	0.175	−0.002	−0.040	0.047
	(0.158)	(0.008)	(0.127)	(0.102)
High ed/Low score x distance change x post	0.236*	−0.003	−0.015	0.080
	(0.121)	(0.014)	(0.236)	(0.140)
High ed/High score x distance change x post	0.118	0.023**	0.433**	0.371***
	(0.150)	(0.010)	(0.171)	(0.117)
Observations	317003	317003	317003	317003
Adjusted R^2	0.077	0.182	0.157	0.187
Dep. var. mean, low ed/low score (2007)	9.533	0.399	7.571	7.923
Dep. var. mean, low ed/high score (2007)	9.755	0.571	9.070	9.521
Dep. var. mean, high ed/low score (2007)	10.034	0.535	8.975	9.292
Dep. var. mean, high ed/high score (2007)	9.923	0.704	10.389	10.971

Note: Distance change is measured in 10 kilometer units, with a mean and max of 0.89 and 1.84, respectively, for treated students in the main specification. All specifications include postal code fixed effects, COMIPEMS exam year fixed effects, municipality-by-treatment linear time trends, covariates, and each of these interacted with the education-by-COMIPEMS score group dummy. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the municipality-by-treatment area level in parentheses.