Vehicles, Travel Demand, and Income: Responses to Subways and Bus Rapid Transit in Mexico City*

Danae Hernández-Cortés† Paulina Oliva‡ Christopher Severen§

Current Draft: December 2020

Abstract

PAPER IS PRELIMINARY; PLEASE DO NOT CITE WITHOUT AUTHOR PERMISSION

Keywords: subway, bus rapid transit, congestion, Mexico City

* We thank Erick Guerra, Rhiannon Jerch, Gabriel Kreindler, and Nick Tsivanidis for useful comments. We thank Natalia Volkow for access to INEGI's microdata.

Disclaimer: This paper represents research that is being circulated for discussion purposes. The views expressed in this paper are solely those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. All errors or omissions are the responsibility of the authors.

†University of California, Santa Barbara. Department of Economics.
‡University of Southern California. Department of Economics.
§Federal Reserve Bank of Philadelphia. Research Department.

1 Introduction

Twenty-four of the world's thirty-one megacities are in developing countries, as are many fast-growing emerging cities. These cities share a common phenomenon: private motor vehicle use increases as cities transition from low to middle income. Private vehicles provide valuable mobility services, connecting households to jobs, services, and social networks. Yet, they also impose substantial public costs: potentially severe congestion and air pollution. Governments provide alternative sources of mobility to urban populations through transportation infrastructure; rail transit, buses, and even bicycles. However, there is still much to learn about how mode and network characteristics impact the patterns of adoption as incomes rise.

In this paper we ask how different modes of public transportation interact to meet demands for mobility across the income spectrum. We study how expansions to the subway and bus rapid transit (BRT) networks shift travel behavior in Mexico City, a dynamic, middle income megacity. Between 1994 and 2017, the region added three subway lines to a system with nine existing lines and built a BRT system of six lines to accommodate rapid population growth and reduce congestion. We leverage the simultaneous expansions of they two rapid transit systems (subway and BRT) to study complementarities in modes of transportation as well as differential adoption across the income spectrum.

In order to study how mode choices shift in response to these expansions, we link the three different waves of origin-destination surveys that were collected over three different decades. The resulting data set is a repeated cross-section that can track census tract-level behavior and district-to-district trips over time. Our paper is the first study to compare subway and BRT in a unified setting and to link the three waves of origin-destination surveys using harmonized geographies (we think).

Our findings suggest that, although proximity to rapid transit predominantly substitutes for other forms of public transportation, substitution away from cars occurs when individuals gain access to multiple modes of rapid transit. In addition, our findings suggest that the substitution away from vehicles is concentrated in top tercile of the income distribution. These findings suggest that, as incomes continue to rise, some strategies of rapid transit expansion can limit the growth in private vehicle use.

Importantly, our findings should be interpreted with caution, as they stem from correlations between transportation use and changes in proximity over roughly 10-year periods. If the placement of the stations is endogenous to local trends in usage, or if individuals sort in response to station placement, our findings would not be denoting changes in behavior in response to rapid transit expansion. In future work we plan to address these possible threats to identification by taking advantage of changes in transportation planning over time and by looking for evidence of sorting in response to the rapid transit expansions we study.

Mexico City faces a set of problems common in large, rapidly developing cities: how to enable mobility without hyper-congestion and air pollution. Congestion and high travel times are two common problems of cities in the developing world (Bryan, Glaeser, and Tsivanidis 2019). As a result, many cities in the world have invested in expensive mass transit projects in order to reduce congestion and travel times. Potential benefits of transit interventions include increases in air quality (Gendron-Carrier et al. 2018), decreases in travel time (Tsivanidis 2018), and reduction of informality (Zarate 2020). The impact of these public transit investments on increasing transit use and reducing private vehicle use is ambiguous. Whether congestion can be reduced by the opening or expansion of transit options will depend on individual travel choice and the substitution between transit and private vehicles.

Other studies have analyzed the role of transport policy in reducing congestion in developing countries and have found mixed results. Hanna, Kreindler, and Olken (2017) show that the withdrawal of a high-occupancy vehicle policy in Jakarta increased traffic and commuting times throughout the city. Therefore, high-occupancy policies have the potential to reduce congestion. Gonzalez-Navarro and Turner (2018) find that subway expansions increase subway ridership, suggesting that there is a ridership response to subway expansions. Similarly, Majid, Malik, and Vyborny (2018) shows that the introduction of a transit line increased public transport by 24% in Lahore. However, in contrast to the successful implementation of BRT in Bogotá (Tsivanidis 2018), Gaduh, Gracner, and Rothenberg (2017) show that Jakarta Bus Rapid Transit system neither increased transit ridership nor reduced motor vehicle ownership. Therefore, whether public transport policies can reduce congestion and travel times remains an open question.[1]

Heterogeneity in travel behavior is also an important factor in the adoption of public transit, especially for middle-income cities with stark socioeconomic differences. Zarate (2020) shows that travel behavior differs for formal and informal workers due to the spatial characteristics of Mexico City. As a result, the author finds that transit improvements decrease informality. Majid, Malik, and Vyborny 2018 show significant public transportation ridership increases among highly educated workers after the introduction of a transit line in Lahore. By studying differences in travel behavior across the income spectrum, our paper also shows that the impacts of rapid transit expansions vary substantially with socioeconomic status.

The rest of the paper proceeds as follows. Section 2 describes our setting and data, as well as public transit trends in Mexico City. Section 3 describes our empirical specification and analyzes travel behavior following the introduction of transit stations. Section 4 analyzes the components of observed changes in travel behavior. In Section 5, we sketch a link to the market access literature that provides a summary statistic for the effects of transit as it relates to vehicle adoption.

1. Severen (2018) shows the large transit infrastructure investments are not necessarily beneficial in the medium run if costs are very high and driving ubiquitous.

2 Setting, Data, and Summary Statistics

Mexico City has a large public transportation network: It's metro is second only to New York City in the Western Hemisphere.[2] The first Metro (subway) line opened in 1969.[3] Construction continued over the next two decades, and by 1991 the system included nine subways lines and a light rail. Despite this, the population of Mexico City roughly doubled over the same time period, leading to increased congestion and pollution.

Mexico continued to expand the subway, but population growth continued to tax both roads and the subway itself. Subway lines became increasingly congested, reducing functional capacity. Furthermore, in 1995, Mexico City's primary public bus service (Ruta 100) collapsed. A variety of small, bus-like (and typically private) services, filled the gap. In the mid-2000s, the city began installing a system of Bus Rapid Transit (BRT) lines, branded as Metrobús (the first BRT line was installed in 2005).

BRT lines generally differ from standard buses in that they have dedicated stations and substantial portions of their routes are on dedicated use lanes. In the case of Metrobús, the dedicated lanes are made of hydraulic cement, resulting in smoother and faster rides. These features typically result in a subway-like travel experience with significantly less exposure to congestion than standard buses. Since the beginning of BRT operations, BRT travel prices have been higher than all other travel modes.[4]

[Figure 1 about here.]

2.1 Trends and Summary Statistics

Figure 1 maps the central part of our study area. Outlined areas filled in white are tracts that we observe in all three periods, while those in gray are tracts (from 1990) that are unobserved in at least one period. Coverage in the central part of Mexico City is good.[5] Figure 1 also shows subway and BRT stations. Mexico city already possessed a substantial subway system of nine lines in 1994

2. Ridership is about 60% of the New York City subway, though if BRT is included ridership is closer to 75% of New York City's. Mexico City has 141 miles of subway lines and 195 stations, whereas New York City 232 miles of subway lines and 472 stations. The current BRT system in Mexico City has 78 miles of routes with 283 stations (BRT usually has station density greater than subways).

3. Interestingly, the Metro was designed to be easily navigable without requiring passengers to read Spanish. To achieve this, each station features a unique symbol in addition to the unique color each line receives.

4. For both Metro and BRT, the price of a trip does not depend on the distance travelled; it is uniform regardless of the distance, time, or number of connections taken within system. The price of a BRT has been higher than the subway. Before 2008, the subway cost $2 MXN ($0.10 USD) per trip and BRT cost $4 MXN ($0.20 USD). From 2008 to 2013, the subway cost $3 MXN ($0.15 USD) and BRT cost $ 5 MXN ($0.25 USD). Since 2013, the subway has cost $5 MXN ($0.25 USD), and BRT has cost $6 MXN ($0.30 USD). Fare payment systems were integrated in 2012. Davis 2020 examines the impact of fare changes in Mexico City to estimate the price elasticity of demand for urban transit showing that ridership responds to changes in prices.

5. Coverage in more distant areas is less consistent. Many of these grayed out areas have experienced substantial growth and change over the two decades in our study.

(shown in black); expansions are shown in gold (for subway) and blue (for BRT). Lines planned in the *Plan Maestro 1985-2010* but not build are shown in green with station locations where available (we discuss this more below).

Vehicle use has been increasing—and transit declining—over the last three decades. Table 1 summarizes mode use and average travel time in Mexico City 1994, 2007, and 2017 (we detail the data below; the data cover a constant geographic area and sample). The majority of non-pedestrian travel is by some form transit (including informal buses) in each year. Nonetheless, transit use has declined: 78% of people and 72% of people used some form of transit in 1994, while by 2017 these shares were only 66% of people and 62% of trips. Buses (including informal buses) have experienced the largest decline in usage, while subway use has been relatively stable. Private vehicle use has increased from 22% of people and 23% of trips in 1994 to 28% of people and 29% of trips in 2017. Taxi use has also increased significantly.

[Table 1 about here.]

Average travel time across all modes has increased by 18%, from 48 minutes in 1994 to 57 minutes in 2017. Each mode has seen an increase in travel times as well (except BRT, which was not operating in 1994 and had only one line operational 2007). Even trips using the subway have seen increased travel times, plausibly indicating congestion in the subway system and increased waiting times. Trips by private vehicle are also substantially longer than they were in 1994.[6]

Trips are also complex. At least 19% of trips required 2 or more non-walking modes throughout our sample. In 2017, 19.5% of trips required 2 modes and 1.4% of trips required 3 or more modes. Nearly all of these multi-modal trips take place entirely on transit. That roughly one-fifth of trips require two or more modes suggests that understanding connections between modes is important.

There is substantial heterogeneity in vehicle use across the income spectrum, as shown by Figure 2. The share of people that use a vehicle remains under 20% for the lowest income individuals, but is well over 50% for higher income individuals. This share using a vehicle increased between 1994 and 2007 across all income levels, but appears to have remained flat since 2007. Indeed, the likelihood of private vehicle use may have even declined a bit at the highest real income levels by 2017.

[Figure 2 about here.]

The differences in vehicle adoption across different income levels suggests that transit impacts will be heterogeneous across incomes levels as well. Moreover, as income levels in Mexico City broadly rise, Figure 2 suggests that more people will adopt private vehicles. Yet, the decline in

6. The increases are not only in mean travel times. The 75% percentile for trips that use the subway for some portion has increased from 90 to 110 minutes since 1994. For private vehicles, the 75% percentile increased from 45 to 60.

Taxi trips in 2017 refer to both traditional taxis and taxis called by ride hailing apps. The share of ride hailing app taxi rides of total taxi rides is 9.4% in our sample.

private vehicle use at the upper end of the income spectrum suggests that there may be some hope of reducing vehicle adoption even while maintaining mobility as incomes rise. We explore the role of transit infrastructure along this dimension in the rest of the paper.

2.2 Data

We briefly describe the sources and most important features of our data below; details are in the Appendix.

2.2.1 Travel and public transit use data

The primary source for travel and household data are three waves of a travel survey conducted in 1994, 2007, and 2017. The microdata consist of three main components: personal and household characteristics, residential location, and trip characteristics (including origin and destination, purpose, modes used, travel time, and cost). The surveys cover households that reside in, or who have members that commute into or out of, the Metropolitan Valley of Mexico, and were implemented by Mexico's federal statistical agency (Instituto Nacional de Estadística y Geografía, or INEGI). We refer to these origin-destination surveys by their INEGI-selected acronym, EOD (for Encuesta de Origen y Destino). We exert significant effort to harmonize both geographic and characteristic data across survey waves, which were unfortunately not designed to be compatible. This consists of three steps: consistent sample selection, geonormalization, and characteristic harmonization.

Consistent Sample Selection. The three EOD waves had slightly different designs, which we correct for by dropping observations so as to create a common sampling design. The 1994 EOD only contains information on people who make at least one non-walking trip. The 2007 EOD contains people who do not make any non-walking trips, but does not capture any information about walking-only trips. The 2017 EOD captures information about all non-walking trips and about individuals who make no trips of any kind. Further, the 2017 EOD captures both weekday and weekend trips, while the 1994 and 2007 EODs only report weekday trips.

Our final sample consists of people aged 15 or older who make at least one non-walking trip on a weekday, and their non-walking trips that occur during the sample weekday. We also exclude people for whom we cannot match to the consistent, geonormalized geographies describe next.

Geonormalization. We use two different geographic delineations for spatial analysis in the EOD: census tracts and travel districts.[7] We geonormalize census tracts to their 1990 geographies, and travel districts to their 2017 boundaries; these are the smallest geographical units to which we can consistently assign most households and trips. In the limited number of cases where it is

7. We refer to INEGI's *áreas geoestadísticas básicas* (AGEBs) as tracts, as they are functionally similar to census tracts in the United States. Travel districts are aggregations of AGEBs used for transportation planning and public disclosure.

ambiguous to which tract or district we should assign a household or trip, we areally assign a proportion of the unit to overlaying areas Autor and Dorn 2013. After normalization, there are 2,478 tracts and 194 travel districts that we consistently observe from 1994 to 2017 (there are an additional 383 tracts that we observe only in 2007 and 2017; we use these in some models). Trips can therefore be reported in two sets of geographies: (i) all trips can be described by district pairs, and (ii) trips that begin or end at home can be described by a pair of district and residential tract.

This geonormalization ensures that we can include tract and district fixed effects, enhancing identification by controlling for local, time-invariant unobservables. However, it does mean that we exclude areas on the urban fringe. Given rapid growth and land use change, studying travel behavior in those areas is worthwhile but would require alternative data sources.

Characteristic Harmonization. The characteristics of trips, individuals, and households are not consistently surveyed or enumerated across the three EODs. We carefully combine characteristics into non-overlapping groups that are consistent across the waves. In some cases, this limits the degree of specificity with which we can study a certain subset of observations. Three variables deserve special attention: vehicle ownership, education, and income.

Both vehicle ownership and education questions where either not asked or not recorded in the 1994 EOD. We therefore generally study whether a private vehicle was the primary mode used by a person. In some cases, we subset based on whether there was a vehicle available in the household. When we do this, we proxy the 1994 variable with an indicator for whether anyone in the household used a private automobile. When we subset by education, we only use data from 2007 and 2017. Instead, we generally use measures of predicted income.

The 1994 and 2007 EODs ask questions about income but the 2017 EOD does not. We therefore draw on an auxiliary data source, the 2015 census, and use that data to predict income out of sample in the 2017 EOD. Specifically, we run Mincer-type regressions, augmented with residential tract fixed effects and characteristics from each EOD to generate predicted income for each employed person in each year, and assign an income of zero to unemployed persons.[8] We generally use predicted income in place of actual income so as to be able to use the 2017 EOD data. See the Appendix for additional details.

2.2.2 Transit infrastructure data

We obtain information on transit infrastructure from the Transportation and Mobility Agency in Mexico City. The data comprises shapefiles of all existing metro and BRT lines and stations in Mexico City and their starting date of operation. We linked the public transit lines to trip data using minimum linear distance. Figure A1 shows a summary of all the Metro and BRT openings

8. We also experimented with machine learning techniques, but these did not generally improve fit much relative to the Mincer models.

with respect to the data sources on trip and population data.

We also use historical data on planned metro lines and stations in Mexico City from the *Plan Maestro 1985* (PM85). PM85 details public transit expansion plans for Mexico City from 1985 to 2010 and contains detailed information on planned routes and station locations (as well as an expansion of light rail, 'tren ligero', that would feed into the expanded subway system). Besides delineating the lines and stations, PM85 provides information of expected users and ranked the lines according to cost and expected users. We obtain historical maps and georeferenced the historic stations to build a map of the entire planned transit network.

3 Transit Proximity and Mode Use

We describe empirical observations following changes proximity to rapid transit stations on travel behavior using a difference-in-differences (fixed effects) strategy. *We use rapid transit to refer to both subway (Metro) and BRT (Métrobus) usage*, excluding other modes.[9] When we refer to transit more generally, we include bus or colectivo use along with rapid transit modes. We briefly describe the empirical strategy below before reporting evidence in support of our empirical observations.

3.1 Empirical Strategy

Let $\mathcal{Y}_{it}$ be the set of modes used by person i reports using in a travel diary in survey year t, so that $1[j \in \mathcal{Y}_{it}] = 1$ if i used mode j and 0 else. Let D^j_{ct} be a measure of proximity of mode j to census tract c in year t. Stack D^j_{ct} across j into D_{ct}, to permit cross-effects of mode proximity, and supplement with interactions if needed. The residential location of person i is tract c_i. We model modal choice in a two-way fixed effects framework:

$$1[j \in \mathcal{Y}_{it}] = D'_{c_i t}\beta_j + X'_{it}\gamma_j + \kappa_{jc} + \varsigma_{jt} + e_{ijct} \tag{1}$$

Equation (1) contains two dimensions of fixed effects. Residential tract fixed effects capture observed and unobserved time-invariant characteristics of residential location c that may influence the use of mode j (such as distance to the central business district or elevation). Residential tract fixed effects also control for average income in the tract from 1994 to 2017, partially capturing neighborhood segregation by income class and limiting the degree to which such segregation can confound our estimates.[10] Year fixed effects absorb year-to-year differences in characteristics that

9. There is not commonly agreed upon term than envelops both subway and BRT, as BRT is often viewed as intermediate between subways and traditional buses. In the context of Mexico City, we view BRT as somewhat more similar to subways because of access limited to stations and mostly separated lanes. We explore heterogeneity between the modes.

10. Such controls are likely to be particularly effective in cities with significant topographic features, such as Mexico City Lee and Lin 2018.

affect the city as a whole, such as average congestion levels, national level wealth shocks, or automobile prices. In some specifications, we also include individual covariates X_{it}. Standard errors are clustered by tract.

There is no clear cutoff distance at which proximity to a transit station becomes much more or less useful. We select a parsimonious specification that utilizes variation in the distance from a station, but also defines a maximum distance at which the marginal effect is zero. This normalized measure of proximity for mode j to location c in year t is:

$$D^j_{ct} = \max\left\{\frac{\bar{d} - \min_{s\in\mathcal{S}_{jt}} d(c,s)}{\bar{d}}, 0\right\}$$

where d is the distance operator, $\bar{d}$ is a maximum distance for which an effect can take place, and $\mathcal{S}_{jt}$ is the set of stations served by mode j in year t. For tracts that contain a station for mode j in year t, $D_{jct} = 1$, which then decreases to $D_{jct} = 0$ for tracts more distant that $\bar{d}$. We set $\bar{d}$ to 2km; if treatment is weakly decreasing in distance, setting $\bar{d}$ to the incorrect distance only attenuates estimated coefficients.[11] For our main model studying rapid transit use, we then define $D^{\text{Station}}_{ct} = \max\{D^{\text{Subway}}_{ct}, D^{\text{BRT}}_{ct}\}$.

3.2 Transit Proximity Results

Observation 1. *Increased proximity to a rapid transit station correlates with increased rapid transit use.*

Columns 1 and 2 of Table 2 show results from Equation 1 on whether or not an individual uses any rapid transit (subway or BRT). The treatment variable, D^{Station}_{ct}, measures the distance of the nearest subway of BRT station. Its associated coefficient captures the difference in usage among those living in a tract that contains a subway or BRT station and those who are at least 2km away from the nearest subway or BRT station (scaled continuously for tracts in between 0km and 2km).

Increasing rapid transit proximity correlates with increases in rapid transit use by up to 18.6 percentage points. This coefficient does not change when individual covariates are added. This large take up suggests that rapid transit modes are preferred to other modes. We explore substitution patterns briefly next, and overall changes to travel behavior later.

[Table 2 about here.]

Observation 2. *Increased proximity to a single rapid transit station does not highly correlate with an increase in any transit use, as there is substitution away from other transit modes.*

Columns 3 and 4 of Table 2 indicate that being near a rapid transit station does *not* increase the likelihood of using transit at all. Coefficients are small and insignificant. This suggests substantial

11. Distances are measured from the perimeter of the AGEB to the closest line. In Appendix Figure A2, we report coefficients from Equation (1) using 250m distance bins for each $j \in \{\text{Subway}, \text{BRT}\}$. Results suggest an effective radius of proximity of roughly 2km. This is a bit muddied by the presence multi-modal trips.

substitution from other transit modes. Indeed, columns 5 and 6 show that bus and colectivo usage declines almost in lockstep with rapid transit increases following station openings. It appears that rapid transit expansions primarily poach users from other transit modes. However, there still could be substantial improvements in travel time, access to markets, and quality of travel, as well as differences by socioeconomics status.

Columns 7 and 8 show that rapid transit expansions are not associated with substantial decreases in the extensive margin of automobile use in and of themselves. These coefficients are either insignificant or marginally insignificant, and relatively small. However, as we point out next, this depends on the setting,

Observation 3. *Increased proximity to multiple rapid transit modes correlates with increased transit adoption overall and decreased automobile use.*

Instead of using $D_{ct}^{\text{Station}} = \max\{D_{ct}^{\text{Subway}}, D_{ct}^{\text{BRT}}\}$ as treatment, Table 3 shows correlations with proximity to Subway and BRT separately as well as interacted. Columns 1 and 2 show that increasing proximity to a subway station or to a BRT station is associated with increased transit use, while Columns 4 and 5 show this proximity is associated with much smaller, but significant, declines in automobile usage.

Columns 3 and 6 include measures of proximity to each type of station as well as their interaction. The interaction term captures proximity to multiple modes, representing richer access to locations served by rapid transit than through either mode alone. Note that access to multiple modes of rapid transit does not seem to increase the overall use of rapid transit beyond the sum of effects of access to either mode (Column 3). However, it appears that it is precisely having access to multiple modes of rapid transit what predominantly reduces the extensive use of vehicle (Column 6). In other words, being *Near both* types of transit leads to a greater decline in auto usage than the sum of either being near subway or BRT alone; and the magnitude of the overall impact on vehicle usage comes mainly from users that gained access to both modes. These results suggest that providing access to rapid transit leads to increased adoption of rapid transit, but that providing travelers with a richer set of travel options is necessary to see substantial declines in automobile use.

[Table 3 about here.]

Observation 4. *The relationship between increases in proximity to rapid transit and travel behaviors vary substantially with income.*

We now investigate heterogeneity in the above results according to income. We provide several measures to ensure that our results are robust to variations in income reporting in the EODs (see Section 2). To this end, Table 4 uses three different samples: Columns 1 and 4 use all available observations and, for those who work, divides them into above- and below-median income groups on the basis of predicted income. Columns 2 and 5 only consider those who work, and use predicted income. Columns 3 and 6 probe how good the income predictions are using the subset of workers for which we have actual income (and thus drops the 2017 EOD).

Columns 1 through 3 show that take up of rapid transit in response to better rapid transit access is greater at lower incomes. Column 1 reports roughly similar responses in rapid transit use among people who are not working as for people who work with incomes in the bottom tercile (about 21 percentage points). For middle-tercile earners, take up in response to rapid transit is slightly lower (18 percentage points). For top-tercile earners, however, take up increases by only about 13 percentage points after a station opening. Columns 2 and 3 allow estimating different responses to increases station proximity as a linear function of log income. At the 10th income percentile, the response is 22-24 percentage points, whereas at the 90th income percentile the response in 10-12 percentage points (just under half the effect size at the 10th percentile).

Columns 4 through 6 repeat this exercise but study automobile use instead of rapid transit use. Column 4 suggests an inverted-U relationship between income and the effect of increased transit proximity; low- and high-tercile earners see a decline in vehicle use of about 3 percentage points after a station opening, whereas middle-tercile earners do not see an effect. Given this result, we next investigate non-linearity in income response.

[Table 4 about here.]

Figure 3 flexibly estimates heterogeneous responses to increased transit proximity in six different predicted income bins among workers: the lowest decile, the 10th-25th percentiles, the middle two quartiles, the 75-90th percentiles, and the highest decile. This exercise uses an alternate definition of D_{ct}, defining it is a binary equal to one if a tract is near (<1km) from a station and zero if else. The plots show mean take up in each income bin for $D_{ct} = 0$ as well as the predicted result with $D_{ct} = 1$ to provide a sense of scale. Panels (a)-(e) show predicted use by income bin comparing treated and untreated locations, while Panels (f)-(j) plot the estimated effects (i.e., the differences between lines in (a)-(e)).

Indeed, Figure 3 shows both substantial heterogeneity and non-linearity. Rapid transit use (panel a) is nearly inverse-U shaped, with lower and especially higher incomes using rapid transit less. Near rapid transit stations, use increases by roughly 14 percentage points for those with below median income, decreasing to less than 5 percentage points for the top decile (see panel f). Private vehicle use (panel b) echoes the shape of Figure 2, showing low use (less than 20 percent) at low income levels, gradually increasing at first but then quickly increase until reaching about 60 percent use at high income levels. There is substantially heterogeneity in response to increased rapid transit access: the relationship is insignificant across the the bottom three quartiles of the predicted income distribution (panel g). It grows across the top two bins, reaching about 5 percentage points for top-decile earners.

The use of any transit (panels c and h) is essentially the inverse of private vehicle use; there is no effect across lower incomes, but a substantial positive effect of stations on any transit use of about 5 percentage points for top-decile earners. These panels together suggest different patterns of substitution for higher- and lower-income workers in response to rapid transit: Rapid transit proximity has no effect on overall transit use or vehicle use at lower incomes, as lower-income workers near stations substitute away from buses and colectivos. High-income workers instead substitute from vehicles to transit.

Panels d, e, i, and j show subway and BRT use based on distance to that mode's nearest station. Subway use mirrors rapid transit generally, showing a declining response to proximity above median incomes and especially in the top quartile. By comparison, BRT's adoption curve is much flatter across incomes, though there is still a substantial difference in adoption between the 75th-90th and 90th-100th percentiles. The difference in response decline for those in the top quartile of income suggests that BRT may be less inferior than subway.

[Figure 3 about here.]

[Figure 4 about here.]

3.3 Towards identification

Using distance to rapid transit stations can be problematic if the construction of new stations is correlated with changes in local unobservables. One possible source of endogeneity is transit placement. For example, city planners may be able to accurately anticipate locations with the most rapid growth in demand for transit, and provide transit to those locations. Or, planners could be more inclined towards placing transit in places that have higher economic activity or benefit people with higher political power.[12] Though the use of tract-level fixed effects controls for some of these concerns, we do not yet fully address these endogeneity concerns.

We currently provide two exercises to support our reduced form results. These are based on the *Plan Maestro 1985-2010* (PM85),[13] which we use to divide locations in Mexico City into several types: (i) locations treated before our sample period, (ii) locations with treatment planned

12. The opposite could also be true. In fact, a group of high income residents in Santa Fe, one of the richest neighborhoods in Mexico City, made a significant lobbying effort to avoid a Metrobus line that would connect Santa Fe to the closest Metro station.

13. The PM85 planned a combination of metro expansions and bus and *Tren Ligero* routes to feed the existing network. The PM85 divided the expansion plans in four stages. The first stages involved extending three existing lines by 1988 (adding three stations, one to each line in the existing network), which was successfully completed. The second stage also involved extending two existing metro lines by 1994 which was not completed due to topological constraints in the southern part of the city (Ciudad Universitaria). Part of these expansions were later built as Metrobus in 2005-2006. The next expansion plans involved building new lines. Among them included the construction of Line B (finalized in 1999) which slightly deviated from its original plan due to problems acquiring land to existing landowners and Line 12 (finalized in 2012). Other Metro lines planned by PM85 involved new lines that were eventually built as BRT lines: BRT line 1 (PM85 line 15), BRT line 2 (PM85 line 8), and BRT line 6 (PM85 line 6).

in the PM85 that were eventually treated, (iii) locations with treatment planned in the PM85 that have not (yet) received treatment, (iv) locations with no planned treatment in the PM85 yet that nonetheless have received treatment, and (v) locations with no planned treatment in the PM85 that have not received treatment. Figure 1 shows the stations in 1994 and planned lines/stations by PM85. The reasons for built infrastructure diverging from the PM85 include topological constraints, historical preservation of archaeological sites, and budget constraints due to the 1985 earthquake and Mexican peso Crisis.[14]

First we show that results are relatively comparable across different combinations of planned and actual treatment and control groups. Table 5 shows results from an enriched version of Equation (**??**) that incorporates PM85 status. Of particular interest are the joint hypothesis tests in the lower panel. The first tests the null hypothesis that the coefficients on $D_{ct}^{\text{Station}} \times 1[\text{Planned \& Treated}]$ and $D_{ct}^{\text{Station}} \times 1[\text{Unlanned \& Treated}]$ are identical. The failure to reject this null hypothesis suggests that among places that were eventually treated, the PM85 did not select locations with different trends in travel behavior.

[Table 5 about here.]

The next hypothesis test looks for differential trends over time in places that were *Ever Treated* from places that were *Planned but Not Treated*, conditional on treatment. The failure to reject the null hypothesis of similar trends conditional on actual treatment status suggests that locations schedule to receive transit in the PM85 that did not receive transit are on similar paths as places that eventually received transit. This is most similar to the experiment in Zarate 2020, which finds no evidence of differential pre-trends between the planned but unbuilt Line C and the constructed Line B.

The last test compares trends in places that were not scheduled to receive transit and never did to places that were scheduled to receive transit but never did. The null hypothesis of similar trends would suggest that places that the PM85 selected but that did not receive transit are evolving in similar way as unconnected parts of Mexico City in general. While we fail to reject the null for rapid transit use, we reject the null for transit use in general and private vehicle use. This means that, in general, the average place without transit in Mexico may not form a good counterfactual for locations that were planned to receive transit. We plan to use these results to inform identification of the prior effects.

We also examine pre-trends. Figure A4 show the differences in neighborhood characteristics using infrastructure variables such as percent of households with access to drinking water, elec-

14. The PM85 had information on the number of expected passengers, topological characteristics such as steepness and rough terrain characteristics such as existing water infrastructure. However, other limitations to the building of these lines were found as the construction started, according to interviews with Mexico City metro authorities. For example, current metro Line 12 had to change the route due to archaeological conservation, current BRT Line 1 had to be built as BRT instead of metro because of topological constraints, and metro Line B had to change route due to failure to buy communal land to landowners.

tricity, and drainage and population characteristics such as total households, total population over 15, and literacy rates. Figure A4 shows the differences between the three groups using the Plan Maestro 1985 before and after 1994. There is no statistical difference between these groups for the variables we examined. Table 5 shows the results of interacting the distance to each modality of rapid transit with an indicator for each of the groups using the Plan Maestro 1985 classification. These specifications compare groups (1) and (2) to group (3) and are restricted to places that experienced a rapid transit expansion after 1985.

3.4 Interpretation

Transportation infrastructure can induce sorting. We check whether sorting on observables is a first order issue. Table tab:sortingsobs shows correlations from Equation (1), but where the outcomes are individual characteristics. Insignificant results would suggest that sorting on observables are not a first order concern. Of course, sorting on unobservables could still be a factor.

Indeed, the results in Table tab:sortingsobs are for the most part insignificant and precisely estimated. People are similar across locations near and far from rapid transit stations in terms of percent working, percent female, percent student, and percents having completed high school and also college. However, on average people near new transit stations are about one year older than those that are distant. There may be some difference in income following treatment: our predicted measure of income falls by a marginally significant 3.8 log points. However, interpretation of this is tricky. Given that formality increases in a subset of locations that receive transit (Zarate 2020), the negative effect could suggests some selection of lower-income households toward new transit stations. However, given that Zarate (2020) finds little evidence of changes in household composition after Line B construction, this could also represent entry into the labor force of lower-earning workers.[15]

[Table 6 about here.]

4 Components of Change

Rapid transit, being rapid, should improve travel speeds for trips that are accessible by the rapid modes. Higher speeds should also enable people to travel farther, opening new work and consumption opportunities. To investigate these margins, we turn to trip-level data. Our panel data on trips data allow us to very flexibly control for unobservable factors that might confound the effect of increased rapid transit proximity on travel time.[16]

15. Gonzalez-Navarro and Quintana-Domeque (2016) find little evidence of changing household composition following road construction in urban Mexico.

16. Many studies of urban mobility use simulated travel times (e.g., from Google Maps). The ability to examine how travel times respond to transit within a city more over a period of more than twenty years is relatively unique feature of our data.

How much do travel times change, and how much is due to changes in travel speed versus mode versus destination? To motivate the following analysis, define average log travel time from a given origin is just the weighted average of travel times from that origin across modes: $\ln(T_o) \equiv \sum_m \ln(T_{m|o})\pi_{m|o}$.[17] To simplify, consider the travel from a location to consist of two features: distance traveled and speed. A physical identify gives: $T_{m|o} = \text{Distance}_{m|o}/\text{Speed}_{m|o}$. We are interested in decomposing the components of travel time change in response to mass transit station openings. Thus, differentiating with respect to station proximity gives:

$$\frac{d\ln(T_o)}{dD_o} = \sum_m \left(\frac{d\ln(T_{m|o})}{dD_o}\pi_{m|o} + \ln(T_{m|o})\frac{d\pi_{m|o}}{dD_o} \right) \tag{2}$$

$$\frac{d\ln(T_{m|o})}{dD_o} = \frac{d\ln(\text{Distance}_{m|o})}{dD_o} - \frac{d\ln(\text{Speed}_{m|o})}{dD_o} \tag{3}$$

The first equation combines changes and mode choice with changes in travel distance and speed. The analysis in Section 3 isolates changes in mode choice. We now turn to trip level data and analyze the other features.

We estimate high-dimension fixed effects models of how travel time changes in response to transit station openings. Specifically, we model variants of:

$$\ln(\text{Time}_{iodt}) = \lambda D_{ot}(\times D_{dt}) + \kappa_{o(t)} + \varsigma_{dt} + \delta_{od} + X'_{iodt}\gamma + e_{iodt} \tag{4}$$

All models include time-of-day-by-year fixed effects (in 15 minute bins) to control for variable congestion.

In our first set of results, we include a single origin-based treatment variable, D_{ot}, and time-invariant origin fixed effects κ_o. Thus, these results capture how much travel times change regardless of the destination. Column 1 of Panel A captures the total change in travel, corresponding to Equation (2). Travel times fall by 10 log points across all modes (i.e., due to changes in mode, destination, and speed). Conditioning on transit use, times do not fall. Columns 3 through 5 condition on mode choice (for at least part of the trip), and so capture both speed and destination changes in Equation (3). Conditional on rapid transit use, travel times fall by more than 25 log points. Bus and colectivo times fall by 4 log points. Interestingly, private vehicle travel times fall by nearly 13 log points, which is most likely due to changes in destinations, though could also reflect some small changes to speed (like time spent finding parking).

[Table 7 about here.]

Panel B investigates changes in speed, $\frac{d\ln(\text{Speed}_{m|o})}{dD_o}$, through the use of origin-destination pair fixed effects δ_{od} (which subsume origin fixed effects). The origin-destination pair fixed effects flexibly control for destination, and ensure that we are comparing travel times for trips that share

17. The true average also includes a variance term; we currently ignore this for simplicity.

starting and ending points. In other words, they capture time-invariant, idiosyncratic factors that make particular routes faster or slower. These models also include destination-by-year fixed effects ς_{dt}, which captures common changes to the destination across all trip origins. Column 1 of Panel B combines speed changes and mode changes, but controls for destination. Transit station proximity reduces travel times across all modes by 4.6 log points conditional on origin-destination pair. That this is smaller in absolute magnitude than in Column 1 of Panel A indicates that people are traveling to locations *less* distant following a station opening nearby. Conditional on using some form of transit and controlling for destination, travel times decrease by 3 log points.

Columns 3 through 5 of Panel B isolate speed changes following station openings, and—by Equation (3)—differences between Panels A and B capture the effects of destination changes. Increased origin transit proximity increases travel times between origin-destination pairs by 13 log points for rapid transit riders, 2 log points for bus/colectivo riders, and has no significant effect on travel speeds for private vehicles. For rapid transit and bus/colectivo riders, the effect in Panel A is about twice that in Panel B, indicating that about half the improvement in travel time following station openings is due to changes in destinations and half is due to changes in travel speed. For private vehicles, the whole effect is driven by changes in destination (specifically, following a station opening, private vehicle users either stop traveling across all trips or disproportionately switch modes to transit for longer trips).

Panel C studies changes in travel time due to increases in rapid transit connectivity at *both ends* of a trip. It does this by adding origin-by-year fixed effects to the models in Panel B. Because the origin-by-year fixed effects absorb the treatment variable, D_{ot}, treatment is itself interacted in these specifications: $D_{ot}D_{dt}$, capturing the effect of transit proximity at both ends of the trip. The overall reduction in travel time for well-connected origin-destination pairs is 7.5 log points, and is 13.1 log points conditional on any transit use. Columns 3 through 5 again isolate speed improvements for specific modes for these well-connected pairs. Rapid transit speeds improve by 7.2 log points, and trips that use a bus/colectivo speed up by nearly 11 log points. There is also a reduction in private vehicle travel times, pointing to a possible alleviation of traffic congestion between well connected locales.

5 Model Sketch

Proximity to a transit station influences travel behavior in a variety of ways. We turn to a market access approach to better capture the complex ways in which rapid transit expansions in Mexico City have altered the relative advantage of a car to other modes in accessing work and consumption locations. We develop market access terms for residential locations with and without cars, and use these to enrich the structure of Equation (1).

Broadly, this consists of three steps. First, we estimate a model of transit mode choice that conditions on origin and destination locations, similar to Tsivanidis (2018). Given the high share of multimodal trips, we adapt the model of multiple discreteness in Gentzkow (2007) to our setting, allowing agents to combine modes to accomplish a trip. We also allow model terms to vary by (exogenous) agent type.[18]

Second, we combine the results from the first step with origin-destination trip data on commutes (work trips) and consumption trips (all other trips). The trip data are well-modeled by gravity, so we use a gravity estimator to recover origin and destination fixed effects. We intend to show that these fixed effects succinctly capture the relevant features of locations needed to understand travel behavior in the city. We combine the estimated fixed effects with the mode choice results to create market access terms that summarize access to work and consumption opportunities both with and without a car. These terms summarize how useful a car is for travel relative to no car.

Finally, we incorporate the market access terms into Equation (1) as follows:

$$1[\text{Car} \in \mathcal{Y}_{it}] = h(\Phi_{ct}^{\ell,\text{Car}}(g), \Phi_{ct}^{\ell,\text{No Car}}(g)) + X_{it}'\gamma + \kappa_c + \varsigma_{gt} + e_{ict} \tag{5}$$

where g indexes groups and Φ^{ℓ} are the market access terms. We allow the market access terms to take two flavors ℓ: *Workplace Access* and *Consumption Access*. Equation (5) includes tract fixed effects, which capture local time-invariant confounds to market access (e.g., limited parking access, measurement error in our spatial data). Further, group-time fixed effects capture systematic trends that shift car adoption across different groups.

18. We further deviate from Tsivanidis (2018) by allowing individuals of different types to have different preferences across transit modes (in the notation below, the d_{ij} terms vary across g).

References

Autor, David, and David Dorn. 2013. "The growth of low-skill service jobs and the polarization of the US labor market." *American Economic Review* 103 (5): 1553–97.

Bryan, Gharad, Edward Glaeser, and Nick Tsivanidis. 2019. *Cities in the Developing World.* Technical report. National Bureau of Economic Research.

Davis, Lucas. 2020. "Estimating the Price Elasticity of Demand for Subways: Evidence from Mexico." *NBER Working Paper 28244.*

Gaduh, Arya, Tadeja Gracner, and Alexander D Rothenberg. 2017. *Improving mobility in developing country cities: Evaluating bus rapid transit and other policies in Jakarta.*

Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew A Turner. 2018. *Subways and urban air pollution.* Technical report. National Bureau of Economic Research.

Gentzkow, Matthew. 2007. "Valuing new goods in a model with complementarity: Online newspapers." *American Economic Review* 97 (3): 713–744.

Gonzalez-Navarro, Marco, and Climent Quintana-Domeque. 2016. "Paving streets for the poor: Experimental analysis of infrastructure effects." *Review of Economics and Statistics* 98 (2): 254–267.

Gonzalez-Navarro, Marco, and Matthew A Turner. 2018. "Subways and urban growth: Evidence from earth." *Journal of Urban Economics* 108:85–106.

Hanna, Rema, Gabriel Kreindler, and Benjamin A Olken. 2017. "Citywide effects of high-occupancy vehicle restrictions: Evidence from "three-in-one" in Jakarta." *Science* 357 (6346): 89–93.

Lee, Sanghoon, and Jeffrey Lin. 2018. "Natural Amenities, Neighbourhood Dynamics, and Persistence in the Spatial Distribution of Income." *Review of Economic Studies* 85 (1): 663–694.

Majid, Hadia, Ammar Malik, and Kate Vyborny. 2018. "Infrastructure investments, public transport use and sustainability: Evidence from Lahore, Pakistan."

Severen, Christopher. 2018. "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification."

Tsivanidis, Nick. 2018. "The Aggregate and Distributional Effects of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio."

Zarate, David R. 2020. *Factor allocation, informality and transit improvements: evidence from Mexico.* Technical report. Mimeo.

Figure 1: Map of subway and BRT lines in Greater Mexico City (1994 AGEB geography).

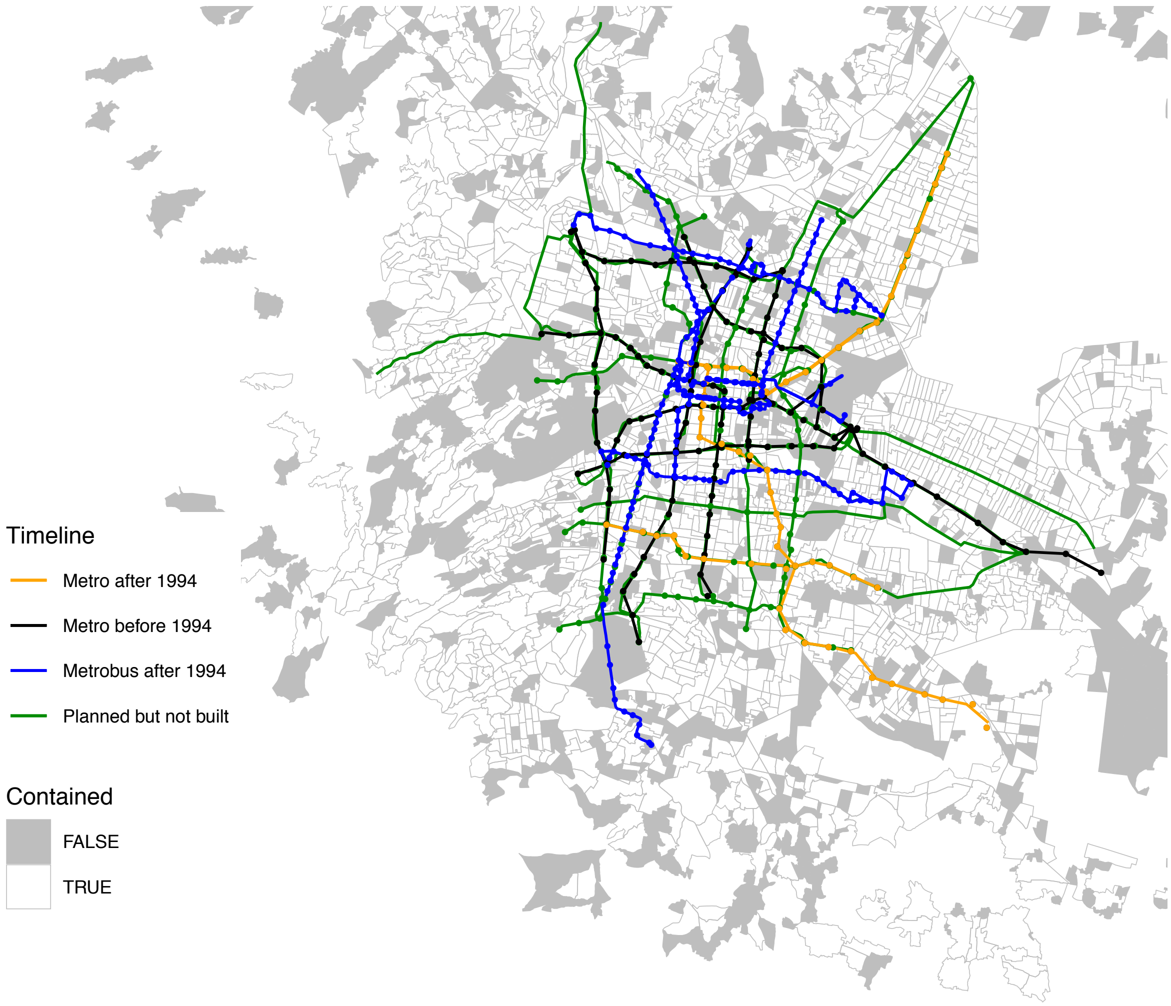


Figure shows the map of our study area along with metro and metrobus stations by year of construction. Outlined areas filled in white are tracts obseved throughout three periods. Shadowed tracts are unobserved in at least one period.

Figure 2: Vehicle use by real income over time.

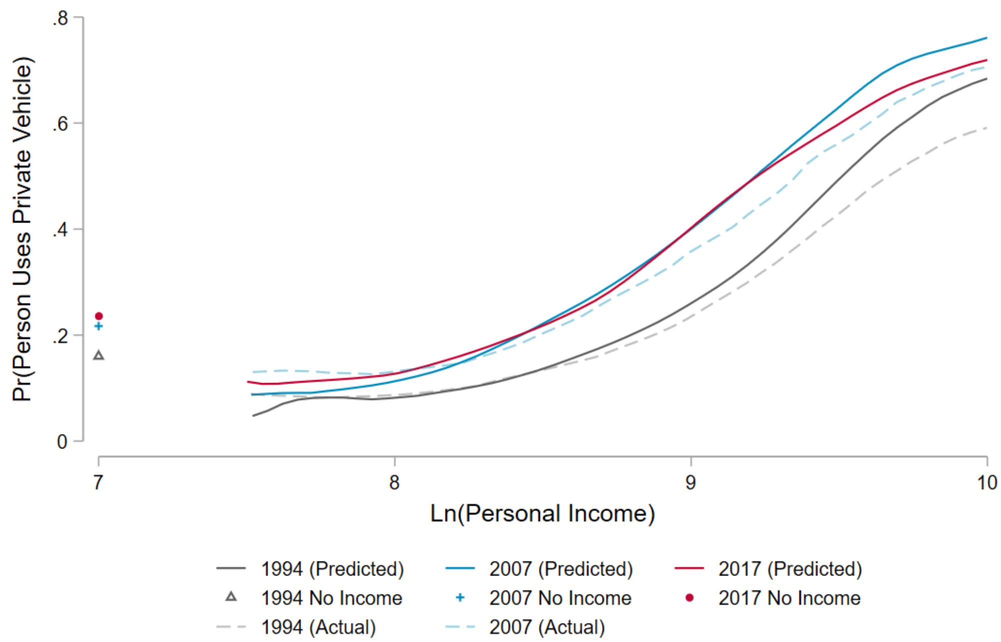
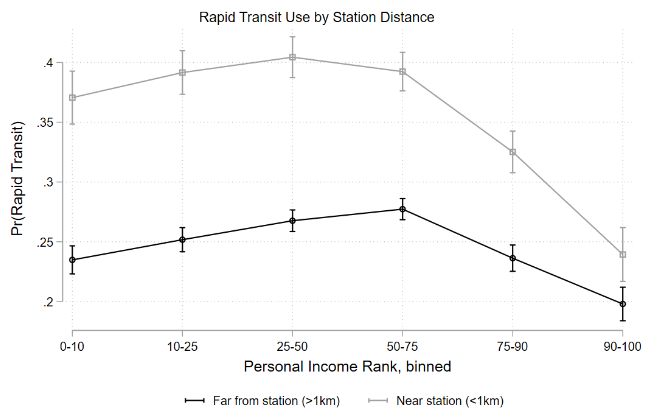
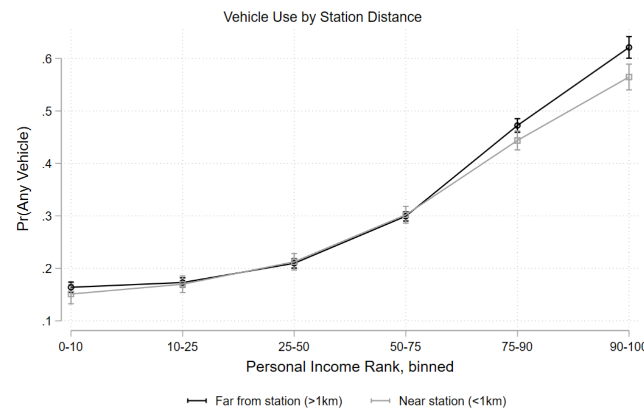


Figure 3: Transit proximity across income bins.

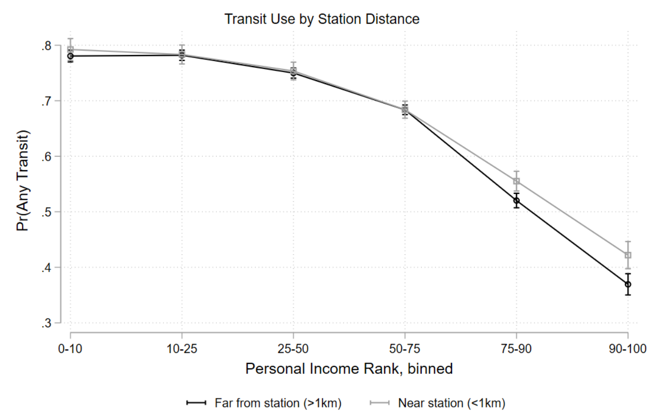
(a) Rapid transit



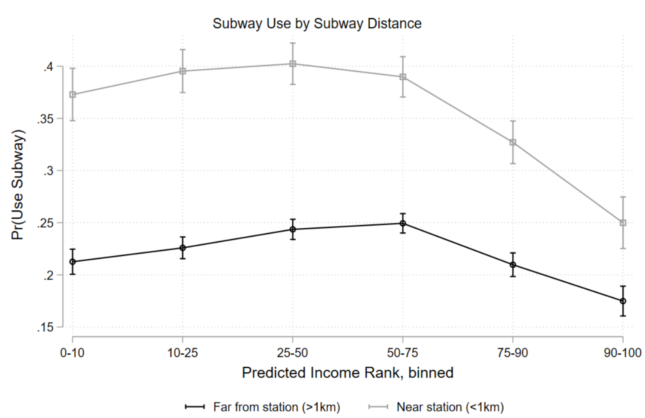
(b) Private vehicle



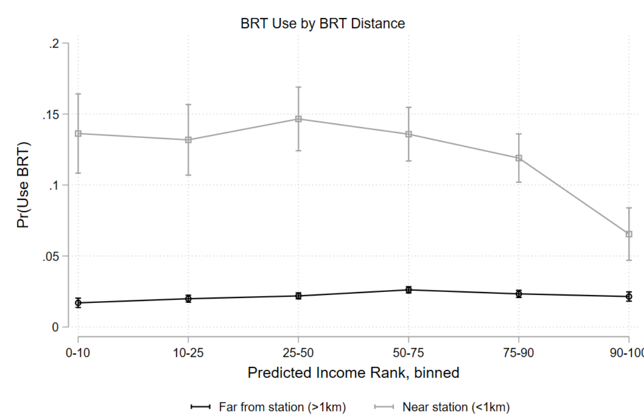
(c) Any transit



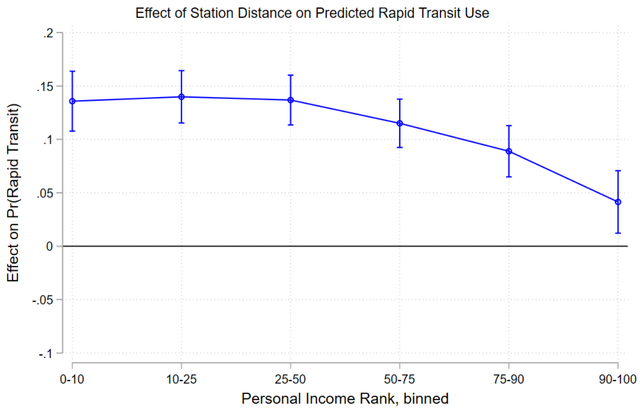
(d) Subway



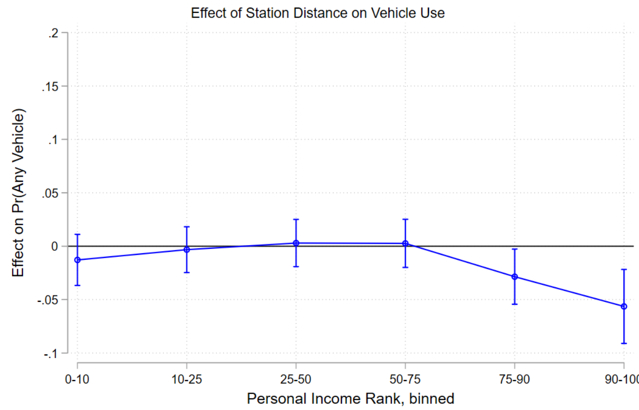
(e) BRT



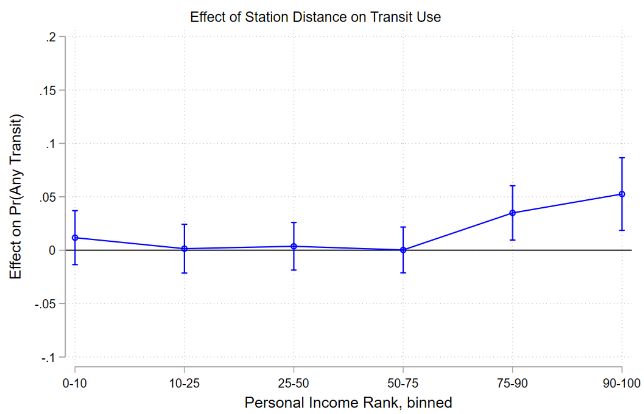
Continued on next page . . .



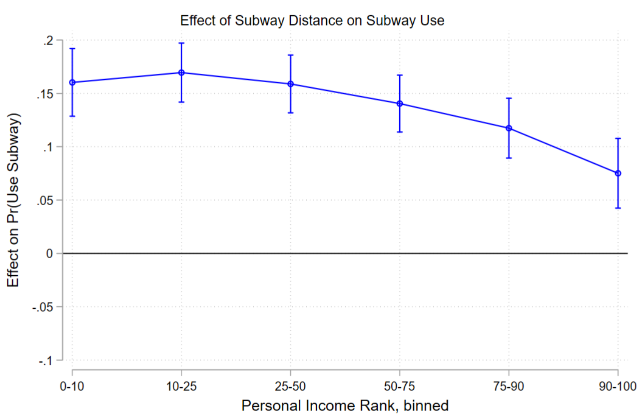
(f) Rapid transit



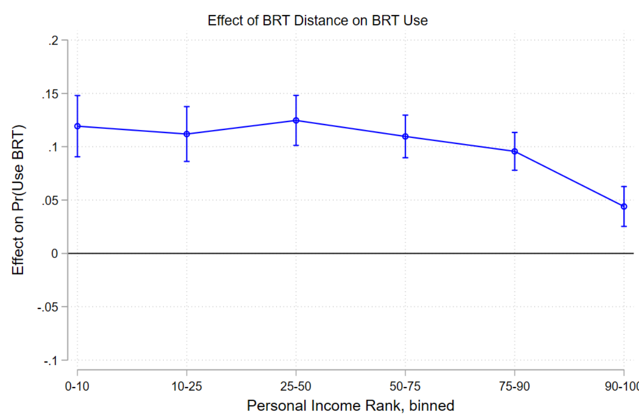
(g) Private vehicle



(h) Any transit



(i) Subway



(j) BRT

Figure shows estimated β from Equation (1) across several modes. The vertical axis denotes the probability of using each mode. The horizontal axis shows predicted personal income rank. Sample only include employed persons. Panels (a)-(c) use D_{ct}^{Station}, whereas Panel (d) uses D_{ct}^{Subway} and Panel (e) uses D_{ct}^{BRT}.

Table 1: Modal choice in 1994, 2007, and 2017.

	1994			2007			2017		
	% People	% Trips	Ave. Time	% People	% Trips	Ave. Time	% People	% Trips	Ave. Time
Panel A. Mode Use Summary and Travel Time									
Transit	0.777	0.723	53.0	0.705	0.628	61.6	0.658	0.614	66.7
Bus	0.757	0.695	53.5	0.668	0.583	62.4	0.576	0.527	67.8
Metro (subway)	0.251	0.210	72.4	0.260	0.211	81.2	0.244	0.220	84.8
Metrobús (BRT)	-	-	-	0.017	0.012	75.4	0.063	0.054	75.3
Private vehicle	0.222	0.234	37.0	0.275	0.290	43.0	0.279	0.285	46.6
Taxi	0.054	0.037	36.2	0.116	0.080	41.2	0.113	0.084	41.8
All			48.2			53.4			56.7
Panel B. Multi-modal Trips									
Trips with 1 mode		0.809			0.804			0.791	
Trips with 2 modes		0.189			0.187			0.195	
Trips with 3+ modes		0.002			0.009			0.014	
Observations	42,501	106,349		68,558	166,492		73,364	153,590	

Sample includes residents of census tracts observed in all three survey years and the trips of such residents. People may use multiple modes, and a trip may consist of multiple modes, so sums do not add to one. Walking-only and biking-only trips are excluded from this table, and walking and biking do not count as additional modes. Travel times for individual trips top-coded to 240 minutes. Observations weighted by sample weights.

Table 2: Station proximity

	1[Any Rapid Transit]		1[Any Transit]		1[Any Bus/Coll.]		1[Any Vehicle]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
D_{ct}^{Station}	0.186***	0.186***	0.009	0.006	-0.151***	-0.154***	-0.019+	-0.013
	(0.011)	(0.011)	(0.011)	(0.011)	(0.013)	(0.013)	(0.011)	(0.011)
Controls	-	Y	-	Y	-	Y	-	Y
N	184693	184670	184693	184670	184693	184670	184693	184670
Unit of observation	Person	Person	Person	Person	Person	Person	Person	Person

Sample includes residents of census tracts observed in all three survey years. Controls include sex, a quadratic in age, employment status, and income rank. Observations weighted by sample weights. Standard errors clustered by residential tract, + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 3: Proximity to different modes

	1[Any Rapid Transit]			1[Any Vehicle]		
	(1)	(2)	(3)	(4)	(5)	(6)
D_{ct}^{Subway}	0.172***		0.188***	-0.022+		-0.027*
	(0.013)		(0.013)	(0.013)		(0.013)
D_{ct}^{BRT}		0.110***	0.166***		-0.047***	0.002
		(0.011)	(0.021)		(0.010)	(0.019)
$D_{rt}^{Subway} \times D_{rt}^{BRT}$			-0.065*			-0.076**
			(0.027)			(0.024)
Near both			0.288***			-0.101***
			(0.020)			(0.018)
N	184693	184693	184693	184693	184693	184693

Sample includes residents of census tracts observed in all three survey years. Controls are sex and a quadratic in age, and are included in all models. "Near both" reports the sum of all three coefficients in the panel above it. Observations weighted by sample weights. Standard errors clustered by residential tract, + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 4: Transit proximity across income.

	1[Any Rapid Transit]			1[Any Vehicle]		
	(1)	(2)	(3)	(4)	(5)	(6)
$D_{ct}^{Station} \times 1$[Not working]	0.205***			0.007		
	(0.012)			(0.011)		
$D_{ct}^{Station} \times 1$[Bottom third income]	0.204***			-0.032**		
	(0.013)			(0.011)		
$D_{ct}^{Station} \times 1$[Middle third income]	0.184***			-0.012		
	(0.013)			(0.013)		
$D_{ct}^{Station} \times 1$[Top third income]	0.131***			-0.031*		
	(0.013)			(0.013)		
$D_{ct}^{Station}$		0.185***	0.167***		-0.019	-0.027
		(0.013)	(0.020)		(0.013)	(0.022)
$D_{ct}^{Station} \times \ln$(Income)		-0.101***	-0.068***		0.010	0.016+
		(0.011)	(0.008)		(0.013)	(0.008)
1[Not working]	-0.073***			-0.025***		
	(0.005)			(0.004)		
1[Bottom third income]	-0.025***			-0.063***		
	(0.004)			(0.004)		
1[Top third income]	-0.035***			0.182***		
	(0.005)			(0.006)		
ln(Income)		0.017**	-0.010**		0.266***	0.164***
		(0.006)	(0.004)		(0.009)	(0.004)
Elasticty at 10th %ile		0.243***	0.224***		-0.024+	-0.040+
		(0.014)	(0.021)		(0.014)	(0.023)
Elasticty at 90th %ile		0.119***	0.102***		-0.013	-0.013
		(0.015)	(0.021)		(0.017)	(0.024)
Income Measure	Predicted	Predicted	Actual	Predicted	Predicted	Actual
Sample	All	Workers	Workers	All	Workers	Workers
Years	All	All	'94 & '07	All	All	'94 & '07
N	181276	123786	67748	181276	123786	67748

Sample includes residents of census tracts observed in all survey years. In columns 2 and 5, predicted income is used with the sample of workers. In columns 3 and 6, actual income is used with the sample of workers, excluding sample year 2017. In columns 2, 3, 5, and 6, ln(Income) is recentered to be mean-zero in sample. Controls are sex and a quadratic in age, and are included in all models. The elasticity measures report the effect of $D_{ct}^{Station} = 1$ evalulated at the 10th and 90th percentiles of income. Observations weighted by sample weights. Standard errors clustered by residential tract, + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 5: Plan Maestro 1985 Comparison

	1[Rapid Transit]	1[Any Transit]	1[Vehicle]
	(1)	(2)	(3)
$D_{ct}^{Station}$	0.234***	-0.017	-0.005
	(0.038)	(0.037)	(0.035)
$D_{ct}^{Station} \times 1$[Planned & Treated]	-0.100*	-0.022	0.017
	(0.040)	(0.039)	(0.037)
$D_{ct}^{Station} \times 1$[Unplanned & Treated]	-0.090*	0.011	0.018
	(0.041)	(0.040)	(0.038)
1[Year=2007] $\times$ 1[Ever Treated]	0.041***	0.034***	-0.028***
	(0.008)	(0.008)	(0.008)
1[Year=2017] $\times$ 1[Ever Treated]	0.070***	0.057***	-0.050***
	(0.009)	(0.009)	(0.009)
1[Year=2007] $\times$ 1[Planned but Not Treated]	0.030	0.046	-0.061*
	(0.022)	(0.028)	(0.027)
1[Year=2017] $\times$ 1[Planned but Not Treated]	0.051*	0.037	-0.051
	(0.025)	(0.034)	(0.033)
1[Year=2007]	0.004	-0.077***	0.065***
	(0.005)	(0.005)	(0.005)
1[Year=2017]	0.002	-0.125***	0.073***
	(0.005)	(0.005)	(0.005)
$Pr(H_0$(Treatments Identical))	0.657	0.144	0.963
$Pr(H_0$(Year Trends Identical across Control/Treatment))	0.748	0.565	0.349
$Pr(H_0$(Year Trends Identical for Controls))	0.190	0.000	0.000
N	184693	184693	184693

Sample includes residents of census tracts observed in all survey years. Census tracts are divided according to the PM85 status and interacted to distance to the closest station/planned station. The second panel reports the p-values from F-tests of the following null hypotheses: (i) the coefficients on $D_{ct}^{Station} \times 1$[Planned & Treated] and $D_{ct}^{Station} \times 1$[Unplanned & Treated] are equivalent, (ii) that the coefficients on Year $\times$ 1[Ever Treated] and Year $\times$ 1[Planned but Not Treated] are equivalent in 2007 and 2017, and (iii) that the coefficients on Year and Year $\times$ 1[Planned but Not Treated] are equivalent in 2007 and 2017, respectively. Observations weighted by sample weights. Standard errors clustered by residential tract, + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 6: Sorting on observable characteristics.

	ln(Income) (actual)	ln(Income) (pred.)	1[Working]	Age	1[Female]	1[Student]	1[≥HS]	1[≥College]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$D_{ct}^{Station}$	0.029	-0.038+	-0.002	1.116***	0.013	-0.008	-0.018	-0.027
	(0.037)	(0.021)	(0.011)	(0.301)	(0.009)	(0.009)	(0.019)	(0.017)
Sample	Workers	Workers	All	All	All	All	All	All
Years	’94 & ’07	All	All	All	All	All	’07 & ’17	’07 & ’17
N	67748	123786	184670	184693	184693	184670	106252	106252

Sample includes residents of census tracts observed in all three survey years. Observations weighted by sample weights. Standard errors clustered by residential tract, + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 7: Travel time

	All Modes (1)	All Transit (2)	Rapid Transit (3)	Bus & Collect. (4)	Private Vehicle (5)
Travel Time					
Panel A. Origin FEs					
$D_{ct}^{\text{Station, Origin}}$	-0.100***	-0.024	-0.255***	-0.044*	-0.127**
	(0.027)	(0.018)	(0.025)	(0.019)	(0.043)
Panel B. Origin-Destination Pair & Destination-by-Year FEs					
$D_{ct}^{\text{Station, Origin}}$	-0.046***	-0.030**	-0.134***	-0.021*	-0.019
	(0.012)	(0.011)	(0.019)	(0.010)	(0.019)
Panel C. Panel B + Origin-by-Year FEs					
$D_{ct}^{\text{Station, Origin}} \times D_{ct}^{\text{Station, Destination}}$	-0.075***	-0.131***	-0.072**	-0.106***	-0.049*
	(0.011)	(0.014)	(0.026)	(0.014)	(0.022)
Mean Travel Time (minutes)	51.3	58.4	75.9	58.8	39.8
N	403952	261522	91015	238310	104799
Number of Modes					
Panel D. Number of Modes (by trip, with orig.-dest.-pair FEs)					
$D_{ct}^{\text{Station, Origin}} \times D_{ct}^{\text{Station, Dest.}}$	0.02+	0.00	-0.22***	0.10***	0.02**
	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)
Mean Number of Modes	1.20	1.30	1.82	1.31	1.01
N	391969	261480	91058	238357	104819

Sample includes trips to and from travel districts observed in all three survey years. All models include origin-destination pair, origin-by-year, and destination-by-year fixed effects, as well as time-of-trip-start-by-year fixed effects (15 minute bins). Observations weighted by sample weights. Standard errors clustered by origin-destination pair, + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Appendix: Additional data details

.0.1 Data Sources

Origin-Destination Surveys

Data come from several sources. Our primary sources for household and travel data are three travel surveys from Greater Mexico City. These surveys cover households in Metropolitan Valley of Mexico in 1994, 2007, and 2017, and were implemented by Mexico's federal statistical agency (Instituto Nacional de Estadística y Geografía, or INEGI). These are:

- *Encuesta de Origen y Destino de los Viajes de los Residentes en el Área Metropolitana de la Ciudad de México* (1994)
- *Encuesta de Origen y Destino de los Viajes de los Residentes de la Zone Metropolitana del Valle de México* (2007)
- *Encuesta Origen-Destino en Hogares de la Zone Metropolitana del Valle de México* (2017)

We refer to these as the EOD, or EOD YYYY for a specific year.

We use two different geographic delineations for spatial analysis in the EOD: áreas geoestadísticas básicas (AGEBs) and travel districts (which we just refer to as districts). AGEBs are roughly similar to a census tract in the United States, while districts are aggregations of AGEBs the represent broader areas. Before geonormalization, there are between 2,500 and 5,000 AGEBs and 100 and 200 districts in each EOD. Residential AGEB is not reported in the public use EOD 2017 (only a coarser residential location is given). INEGI performed a confidential merge that links households to residential AGEBs for this project.

Delineations of AGEBs change over time, but generally correspond to contemporaneous INEGI geodata (see section on Geonormalization below). Thus we use INEGI areal shapefiles where possible. However, there are two exceptions. Instead of standard AGEBs, EOD 1994 uses a modified AGEB layer in which very small AGEBs are combined. We received shapefiles for these modified AGEBs from Erick Guerra. Districts are not a standard INEGI geography; we form them from AGEBs when possible. In EOD 2017, however, we do not know origin and destination AGEB (we only know origin and destination districts), so we use a shapefile of 2017 districts included with the data.

Census Data

We draw on microdata from the 2015 Conteo (an intercensal survey) conducted by INEGI in order to predict income in EOD 2017, which lacks an income measure. We use data from both the Distrito Federal and the Estado de México. Residential AGEB is not reported in the public use Conteo 2015 (only a coarser residential location is given). INEGI performed a confidential merge that links households to residential AGEBs for this project.

Wage Prediction

We do not observe income in the 2017 EOD. Nonetheless, we observe income the 1994 and 2007 EODs and the 2015 Conteo. When we refer to income, we generally use a predicted measure of income so as retain comparability across survey waves. We predict income separately in each year for employed persons with a Mincerian regression of log income on a quadratic in age, gender, residential ageb fixed effects, and a selection of variables that depends on availability form year to year (education, rough occupation and/or industry). We allow for separate wage-age profiles by gender. For 1994 and 2007, we predict income on the same sample that we used for estimated. For 2017, we estimate the income model using the 2015 Conteo and then predict out of sample on the same characteristics that are available in the 2017 EOD.[19]

The top panel of Figure A3 compares the distribution of observed real income in 1994 and 2007 with predicted income in 2017. The distributions largely overlap, but naturally predicted income is smoother and appears to exhibit less variance. The bottom panel of Figure A3 shows average log personal income for each worker in each normalized predicted income centile.[20] This panel shows that the prediction does a reasonable job on average. Centiles of income and log income are linearly related except at the tails. The overall level of real income is slightly higher in the 1994 sample.

19. Attempts at using machine learning methods did not greatly improve fit, likely because of a lack of richly detailed employment characteristics.

20. These are normalized so that the mass of not working persons (about 50% of the sample) have a centile of 0.

Infrastructure Data

We draw geographic data on the location of Metro and Metrobus from the Mexico City transportation agency. Additional data was provided by Diego Valle's webiste. Finally, we date station openings using data from the Mexico City's metro and metrobus websites. We order openings according to the following list and it is captured in the figure A1. The spatial distribution of these lines can be seen in figure 1.

- 1994 (May/June): EOD conducted
- 1994 (July): Subway Line 8 opens
- 1999 (July): Subway Line B opens
- 2005 (June): BRT Line 1 opens
- 2007 (January/March): EOD conducted
- 2009-2012: 3 BRT Lines open (2,3,4)
- 2012 (October): Subway Line 12 opens
- 2013 (November: BRT Line 5 opens
- 2016 (January): BRT Line 6 opens
- 2017 (January/March): EOD conducted

Appendix: Additional tables and figures

[Figure 1 about here.]

[Figure 2 about here.]

[Figure 3 about here.]

[Figure 4 about here.]

Figure A1: Timeline of transit openings

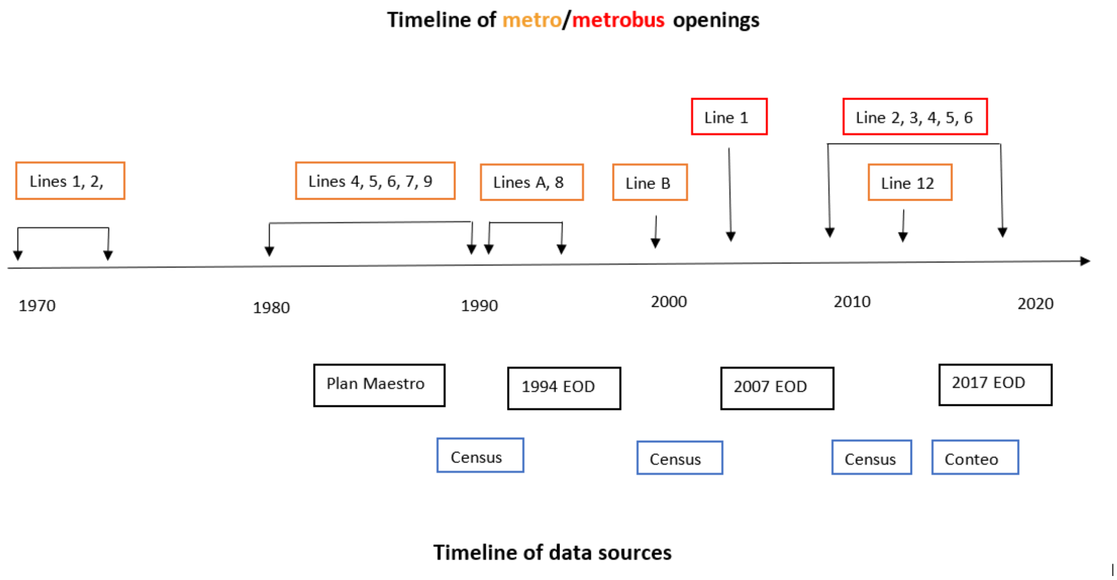
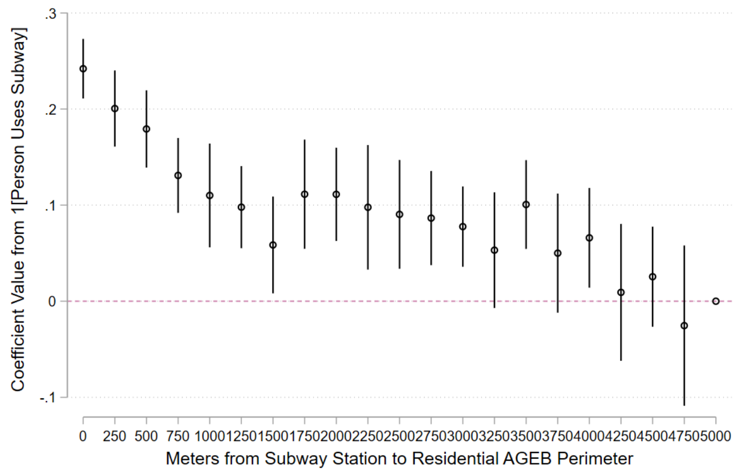
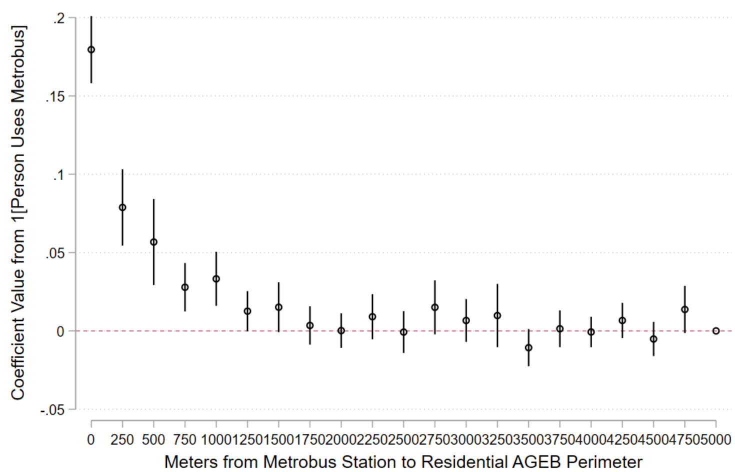


Figure A2: Treatment effect by distance bins and rapid transit type.

(a) Subway



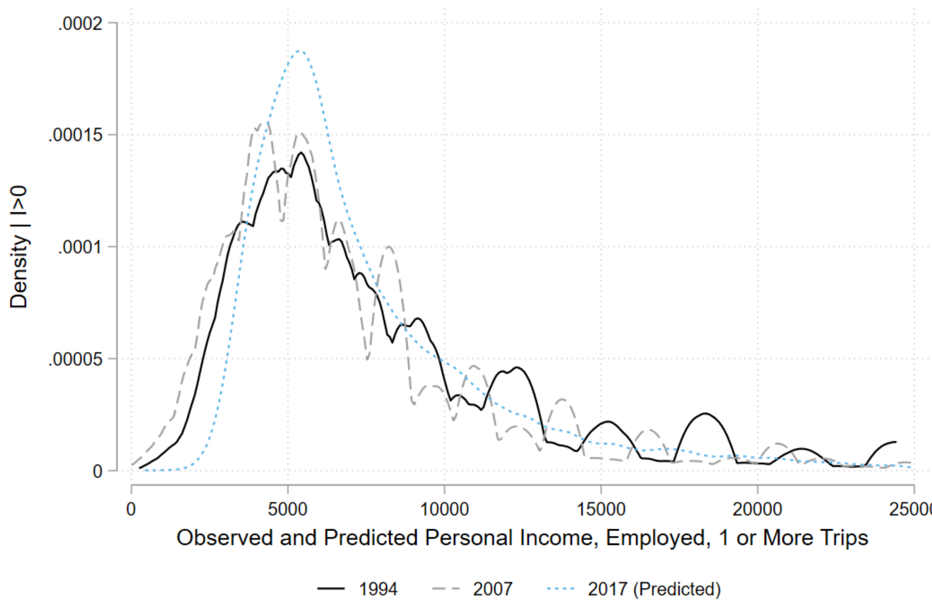
(b) Metrobús



Estimates of Equation (1) with heterogeneity by distance from edge of tract to station. Sample includes residents of census tracts observed in all three survey years. Observations weighted by sample weights. Standard errors clustered by residential tract, + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. , + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Figure A3: Income and Predicted Income.

(a) Comparing actual 1994 and 2007 income with predicted 2017 income.



(b) Comparing prediction incomes by centile if employed across 1994, 2007, and 2015.

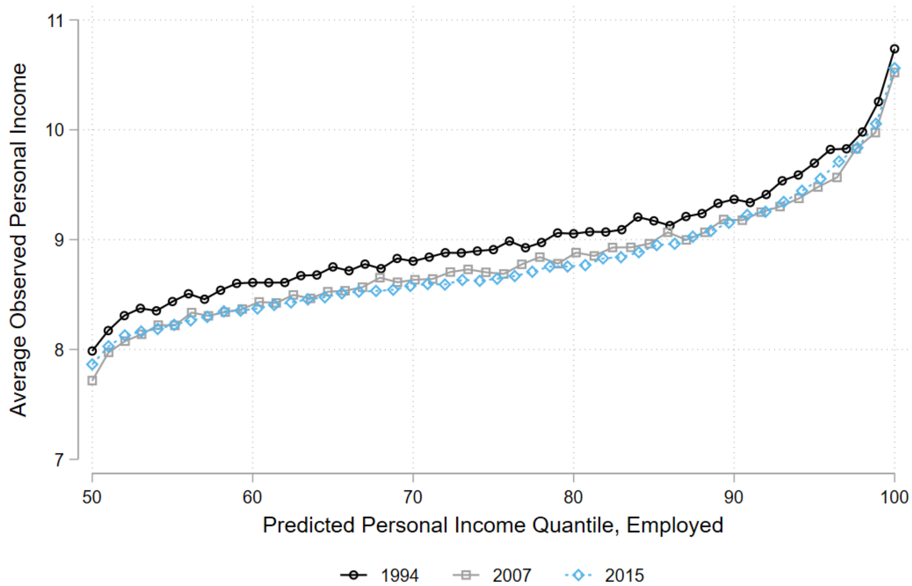
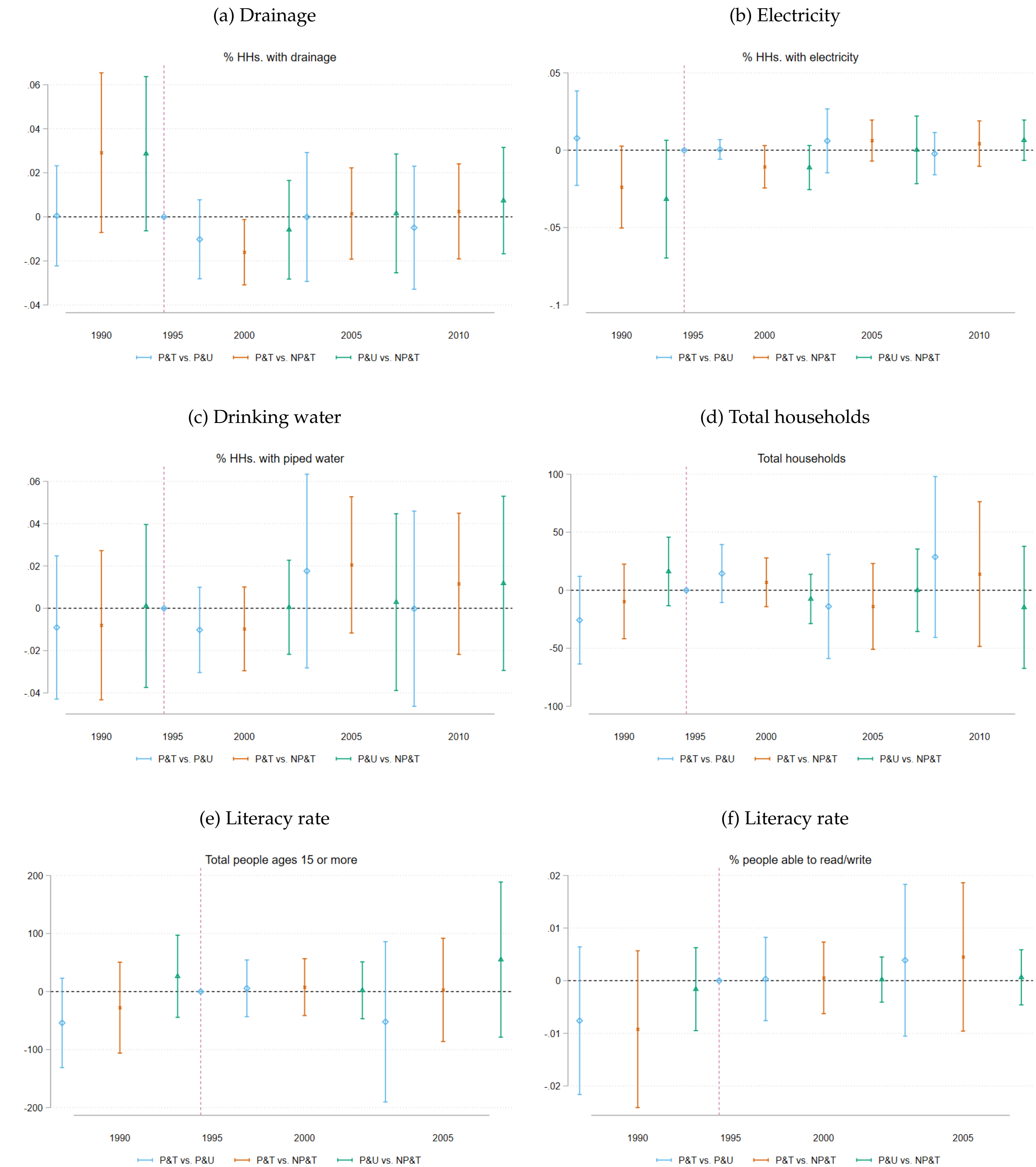


Figure A4: Parallel trends between Plan Maestro 1985 groups



Comparison of Plan Maestro 1985 groups before and after 1994. Sample includes all the AGEBs under the Plan Maestro 1985 classification, + p<0.10, * p<0.05, ** p<0.01, *** p<0.001