Scale Effects of Rapid Transit and Automobile Adoption *

Danae Hernández-Cortés† Paulina Oliva‡ Christopher Severen§

September 20, 2024

PAPER IS PRELIMINARY
DO NOT CITE OR CIRCULATE WITHOUT AUTHORS' PERMISSION

Abstract

We show that scale effects of rapid transit—access to multiple transit lines—is a key determinant of the effect of transit on travel behavior. Households near two or more lines are much less likely to adopt a vehicle than households near only a single line. We study travel behavior in Mexico City, which greatly expanded its transit system during our study window (1994–2017). We build a partial-equilibrium quantitative spatial model to understand the determinants of vehicle adoption. Notably, the model accommodates continuously heterogeneous preferences while retaining attractive aggregation properties. Model results indicate the consumption market access overwhelmingly determines the scale effects of transit. Moreover, because even the highest-income households still value improvements of transit relative to automobile market access, they adjust automobile adoption in response to large transit expansions.

1 Introduction

Investment in public transit infrastructure can reduce congestion, decrease local pollutants and greenhouse gases, and connect people to opportunity. However, there are substantial disparities in the effects of rapid transit infrastructure, particularly with respect to automobile use (Baum-Snow, Kahn, and Voith 2005). Understanding the determinants of these disparities is vital for transit to be an effective policy lever, particularly because infrastructure determines land use in the longer run (Baum-Snow 2007; Brooks and Lutz 2019; Bryan, Glaeser, and Tsivanidis 2020).

We propose that the scale of transit access is key for determining whether it provides a broadly palatable alternative to automobile use. We test for scale effects using large rapid public transit

*We thank Natalia Volkow for access to INEGI microdata. Qitong Wang, Zoey (Zhaoyang) Li, and Nathan Schor have provided valuable research assistance.

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.

†Arizona State University, School of Sustainability and NBER.
‡University of Southern California, Department of Economics and NBER.
§Federal Reserve Bank of Philadelphia, Research Department.

expansions in Mexico City that provide rich variation in transit density. Between 1994 and 2017, Mexico City built three new subway lines and a six-line BRT system. These expansions came on the heels of an already large rapid transit network. The first Metro (subway) line opened in 1969, and just before 1994 (the start of our study period) the system already included nine subway lines. The existing network provides variation across locations in the level of pre-existing transit connectivity, allowing us to explore the marginal effect of expansions at different density levels. In order to explore how these expansions shaped vehicle and public transit adoption, we harmonize (for the first time) three rounds of a detailed travel survey, the Encuesta Origen Destino (EOD), which was conducted in 1994, 2007 and 2017. These surveys contain detailed information both on commuting and non-commuting trips—including modal split, travel time, and distance. In addition, the data include rich sociodemographic information, which allows us to explore heterogeneity of responses. The geographic harmonization of the data implies that we can use small-area residential fixed effects in our research designs, such that estimates of transit's effect are free of potential time-invariant confounders.[1]

Access to a single transit line near a household's residence leads to at most a mild reduction in automobile adoption of 3.0pp. However, there are significant and substantially larger effects of additional lines. Receiving a second line transit decreases automobile adoption by an additional 11pp, and a third line further decreases adoption by another 11pp. Taken together, access to three or more rapid transit lines decreases automobile adoption by a striking 25pp, which is about half of the baseline rate. These effects are substantially larger than in many existing studies that mask heterogeneous scale effects by defining treatment based only on distance to the the nearest transit station (e.g., Baum-Snow and Kahn 2000; Baum-Snow, Kahn, and Voith 2005; Kahn 2007).

We then explore the potential mechanisms behind these scale (or density) effects. The scale effects in the reduced form results could be driven by positive network spillovers increasing access to labor or consumption markets—every additional rapid transit line that serves a given location could be leveraging an increasingly (locally) dense network. For additional lines to be relevant, households should value access multiple destinations.[2] Because we have trip data for both consumption and labor purposes, we can explore whether the scale effects we find in the reduced form are reflected in their corresponding measures of market access at the household level. Alternatively, scale effects could reflect a role for additional transit lines in luring individuals with increasingly higher value of time. Income is a primary determinant of the value of time for travel decisions (Deacon and Sonstelie 1985; Glaeser, Kahn, and Rappaport 2008; McFadden 1974).

We develop a structural model of automobile adoption to study whether the observed scale

1. We also explore results of alternative research designs that relax the parallel trends assumption needed for identification in our baseline design. Estimates are broadly consistent across designs.

2. For these effects to be relevant to the transit adoption decision, the value of different destinations—either for work or consumption access—should be somewhat homogeneously distributed across space (i.e., not too clustered). Otherwise, connecting residential locations to the dominant consumption or labor market could be easily achieved by access to a single transit line. We show that this is the case, particularly for consumption travel.

effects operate through labor market access, consumption market access, or heterogeneous preferences for transit by income. The model endogenizes the choice of trip destinations and allows multi-modal trips.[3] Explicitly incorporating multi-modal choices allows us to capture the complementarities across modes within the network. The expected value of utility maximized over multi-modal choice sets—which do not include a vehicle option for non-vehicle owners—captures accessibility to each location either with or without a vehicle. Transit expansions reduce transit travel times to some destinations, increasing accessibility. Trip destinations are of two types, consumption and work; we model accessibility to each type of location differently. For work locations, we assume working households are initially endowed with the workplace where we observe them working in the survey. Thus, accessibility to that particular workplace matters for their vehicle adoption decision. However, their decision also reflects the possibility that their workplace may change in the future. This hybrid structure, which is consistent with costly labor force transitions, allows for flexibility in the order of car ownership and workplace decisions.[4] In contrast, our model assumes that consumption location decisions are always made after the car ownership decision, and so the expected value over destinations yields consumption market access terms that differ by vehicle ownership.

Because people's choices over destinations reveal their value, the model captures the heterogeneity in connectivity value across lines that seem to play an important role in transit adoption in our reduced form analysis. At the top of the decision tree, the choice to adopt an automobile (conditional on residence) reflects the relative market access for both work and consumption provided by vehicles compared to transit.[5] At every decision step, the model accommodates rich heterogeneity in preferences using a latent-class formulation in which income determines class assignment probabilities.

Model estimates allow us to differentiate scale effects of transit expansions on work market access, consumption market access, and heterogeneous preferences by income (including value of time). We find large scale effects on consumption market access, but not on work market access. This finding suggests that additional lines connect households to sparsely distributed consumption points.[6] The model reports heterogeneous valuations of relative market access by income. While households value work market access less as incomes rise, they value consumption market access more. Notably, the valuations of market access are significantly above zero even in the upper tail of income, suggesting that sufficiently dense transit can be an effective policy tool that can shift automobile adoption even for those with high incomes.

3. Multi-modal trips are those that use multiple modes. For example, individuals can travel by taking a bus that connects them to a BRT line or by taking a BRT line and then a *pesero* (an informal transit mode) to their final destination.

4. This approach yields an indirect utility of work access that is a weighted average of access to the existing workplace and other potential workplaces.

5. We also demonstrate how to endogenize residential choice in the Appendix.

6. We verify that consumption destinations are sparser than workplaces by mapping workplace and consumption destination fixed effects (which measure value of destinations). Consistent with this explanation, we find that destinations with large consumption value are more disperse than destinations with large work value (productivity).

Our paper offers several contributions. First, we provide a simple and powerful way to measure scale effects of transit—the number of unique transit lines accessible at residence. The existent literature in economics generally measures whether there is access to any rapid transit at all (Baum-Snow and Kahn 2000; Baum-Snow, Kahn, and Voith 2005; Billings 2011; Chen and Whalley 2012; Heilmann 2018; Glaeser, Kahn, and Rappaport 2008; Kahn 2007; McMillen and McDonald 2004). Credible identification typically requires studying changes in transit infrastructure and many transit expansions take the form of new systems, which are typically sparse. Only occasionally large, existing systems that receive extensions are the object of study (Gibbons and Machin 2005; Gupta, Van Nieuwerburgh, and Kontokosta 2022). Yet it is in such places that most variation in transit density can be found. We show that density of the network matters for the response to transit expansions. Moreover, estimates suggest that failing to account for transit density substantially understates the effect of public transit on travel behavior.

We also contribute by developing a model that flexibly measures the value of an automobile and uses transit expansions over time to credibly identify the value of market access to the automobile adoption decision.[7] Recent literature typically focuses on the choice of workplace and does not endogenize automobile adoption (Ahlfeldt et al. 2015; Almagro et al. 2024; Barwick et al. 2021; Conwell, Eckert, and Mobarak 2023; Gaduh, Gračner, and Rothenberg 2022; Severen 2023). Our findings buttress recent papers that have found an important role of consumption market access in urban settings (Lee and Tan 2024; Miyauchi, Nakajima, and Redding 2021). Existing research that endogenizes vehicle choice makes the strong assumption that the parameter that governs the elasticity of automobile adoption also governs residential and/or workplace choice (Rothenberg and Du 2023; Tsivanidis 2023). In contrast, we separate the vehicle choice from other decisions. This has two important advantages. First, it permits identifying the elasticity of automobile adoption solely from expansions in the transit system rather than confounding with factors that drive residential and workplace choice. Second, our model leaves room for other motives of automobile adoption (e.g., conspicuous consumption) which may vary by income.

We also offer a new modeling approach to flexibly capture heterogeneity in quantitative spatial models. Existing approaches to heterogeneity in this literature utilize discrete classes (Diamond 2016; Fajgelbaum and Gaubert 2020; Tsivanidis 2023) or allow only a monotonic relationship between income and the marginal value of time (Almagro et al. 2024). In contrast, we develop a latent class representation (e.g., Lind and Ramondo 2023) of continuous preference heterogeneity that aggregates into discrete groups that retain conformability with the tools typically used with

7. Unlike papers that rely on flow data (e.g., Gupta, Van Nieuwerburgh, and Kontokosta 2022; Miyauchi, Nakajima, and Redding 2021), we focus our analysis on automobile adoption. Automobile adoption is a more holistic and longer-run measure of the utility of transit relative non-transit travel because it reflects the many different motivations and destinations that a household may have. Davis (2008) shows that driving restrictions have unintended consequences for vehicle decisions, and Severen and Van Benthem (2022) reveals that vehicle access may reflect deeply held preference heterogeneity.

quantitative models (Redding 2023).[8] This approach offers a flexible alternative the distributional assumptions that often underly parameter heterogeneity in choice models (e.g., Berry, Levinsohn, and Pakes 1995). The flexibility of this approach is apparent in our results, which reveal non-linearities that would not be detectable with two discrete groups.

Finally, we contribute to the increasingly important literature on transportation technology adoption in developing and middle-income cities (Balboni 2019; Balboni et al. 2020; Conwell 2023; Kreindler et al. 2023; Oliva 2015; Tsivanidis 2023; Zárate 2023). Relative to this literature, our results highlight the role of access to consumption opportunities in determining automobile adoption and demonstrate the transit interventions can be impactful even at the high end of the income distribution. Understanding the determinants consequences of how technologies (and especially cars) spread is vital to managing their future impact on cities, energy demand, and the environment (Davis and Gertler 2015; Oliva et al. 2020; Zhou and Li 2018).

We next describe our study context, Mexico City, and the data sources we develop. We then discuss our primary reduced form designs in Section 3 and report results in Section 4. To differentiate mechanisms, we develop a quantative model of trip and vehicle choice in Section 5 and detail estimation and results in Section 6.

2 Setting, Data, and Summary Statistics

Mexico City has a large public transportation network: Its metro is second only to New York City in the Western Hemisphere.[9] The first Metro (subway) line opened in 1969, and just before 1994 the system included nine subway lines. The Metro continued to grow, adding three more lines by 2017. In 2005, Mexico City began building a system of Bus Rapid Transit (BRT) lines, called Metrobús. This system consisted of six lines by 2017; further lines have since been added. Table A1 shows the timeline of Metro and Metrobús openings between 1994 and 2017. The Metropolitan Area of Mexico City (which includes Mexico City and some municipalities of the State of Mexico) has more than doubled in population since the first Metro line opened in 1969, leading to increased congestion and pollution.

BRT lines 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 Metrobús operations, Metrobús fares have been higher than Metro

8. Although we do not pursue defining general equilibrium due do data constraints, nothing about our approach the latent-class approach prohibits this.

9. Mexico City has 141 miles of subway lines and 195 stations, whereas New York City has 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 subway).

or regular bus fares.[10]

Figure 1 maps the central part of our study area and shows Metro and Metrobús stations. The nine subway lines built before 1994 are in black; expansions are shown in orange (for the subway) and blue (for BRT). Lines planned in the *Plan Maestro 1985* (discussed later) but not built are shown in green with station locations.

[Figure 1 about here.]

Mexico City is also served by a public bus system, concessioned bus lines, and informal transit, as well as a few other miscellaneous modes of transit trolleybus, a light rail line, and new commuter rail and suspended aerial lift lines). To ease exposition, we use **rapid transit** to refer to subway and BRT. Similarly, we often use **informal transit** or **bus & *colectivo*** to refer to any non-rapid transit.[11,12]

2.1 Data

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

2.1.1 Travel Surveys (EODs)

The primary source for travel and household data is three waves of a travel survey conducted in 1994, 2007, and 2017 (Encuesta de Origen-Destino, EOD). The data 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 have members who commute to, the Metropolitan Area of Mexico City. EOD 1994 and 2007 surveys were run in May and June, whereas EOD 2017 was run in February and March.

The three EOD surveys were not designed to be comparable or compatible. To overcome this limitation, we carefully harmonize geographies and characteristics across survey waves. This

10. Before 2008, Metro's fare was $2 MXN ($0.10 USD) and Metrobú's fare was $4 MXN ($0.20 USD). From 2008 to 2013, Metro's fare was $3 MXN ($0.15 USD) and Metrobú's fare was $ 5 MXN ($0.25 USD). Since 2013, Metro's fare is $5 MXN ($0.25 USD), and Metrobús's fare is $6 MXN ($0.30 USD). For both Metro and Metrobús, the price of a trip does not depend on the distance traveled; it is uniform regardless of the distance, time, or number of connections within system (except for airport service). For the purposes of our structural model, this means that mode-specific fixed effects—which we allow to vary over time in order to reflect changes in service quality—will absorb any price variation over time. Davis (2021) examines the impact of fare changes in Mexico City to estimate the price elasticity of demand for transit, showing that ridership responds to changes in prices.

11. This is a slight abuse as some of the buses are technically formal, however, the travel survey data we use does not consistently distinguish non-rapid bus types nor registered buses from truly informal (unregistered) modes (e.g., *colectivos*, *micros*, *combis*, and *peseros*).

12. These informal transportation modes, generally refer to small vans (*combis*) and buses (*micros* or *peseros*) that cover fixed routes and are operated by private providers. These modes are generally used as a last-mile form of public transit and the routes usually start in Metro stations.

consists of three steps: consistent sample selection, geonormalization, and characteristic harmonization. We briefly summarize these steps:

Consistent Sample Selection. Each EOD wave has a slightly different sampling design. Specifically, EOD 1994 only contains information on people who make at least one non-walking trip.[13] EOD 2007 contains people who do not make any non-walking trips, but does not capture information about walking-only trips. EOD 2017 captures information on all non-walking trips and on individuals who make no trips. EOD 2017 also captures both weekday and weekend trips, while EOD 1994 and EOD 2007 only report weekday trips. After excluding observation with criteria that are not captured in all EOD waves, the combined sample consists of people aged 15 or older who make at least one non-walking trip on a weekday and their non-walking trips.

Geonormalization. Residences and trips are reported using two different spatial scales, census tracts and travel districts, respectively.[14] Public use EOD 2017 data only include travel district location. We instead use a confidential version of EOD 2017 microdata from INEGI; this allows recovering census tract of residence. We geonormalize all census tracts to their 1990 definitions and all travel districts to their 2017 boundaries. These are the smallest geographical units to which we can consistently assign households and trips. When tract or district assignment is ambiguous, we assign a proportion of the unit to the overlaying areas. After normalization, there are about 2,478 tracts and 200 travel districts observed in all three EODs. For trips, we use two sets of geographies: (i) trips that begin or end at home (including most commutes) are assigned to tract$\times$district, (ii) every trip can be assigned to district$\times$district.[15]

Characteristic Harmonization. The characteristics of trips, individuals, and households are not consistently surveyed or enumerated in the three EODs. We carefully reclassify characteristics into a consistent scheme. In some cases, this limits the specificity with which we can study a certain subset of observations. Three of these variables deserve special attention: vehicle ownership, education, and income. Vehicle ownership was not recorded in EOD 1994. Therefore, we sometimes proxy vehicle ownership with household private vehicle use.[16] Education was not recorded

13. We believe that data was collected on people who did not make any trips, but those observations are not in the data we recovered.

14. 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.

15. Geonormalization ensures that we can include tract and district fixed effects, enhancing identification. 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.

16. Importantly, we observe trips for work and non-work purposes for every member of the household above 15 years of age. Thus, it is very likely that if no member of the household used a vehicle in any of the trips observed, the household does not own a vehicle. We check the accuracy of this ownership measure using the survey waves for which we have an explicit vehicle ownership question, and we find that household vehicle use is equivalent to ownership for 81% of households.

in EOD 1994. EOD 1994 and EOD 2007 record personal income but EOD 2017. We therefore draw on auxiliary data to predict income out of sample in EOD 2017. We use confidential 2015 census data from INEGI for this purpose (see Appendix, Section C for details).

2.1.2 Transit infrastructure data

We use GIS data on Metro and Metrobús lines and stations in Mexico City, combined with starting date of operation. We link the transit stations to locations using minimum linear distance to the nearest edge of the location. We also use historical data on planned metro lines and stations in Mexico City from *Plan Maestro 1985* (PM85). PM85 details expansion plans for public transit in Mexico City from 1985 to 2010 and contains information on planned routes and station locations. We georeference PM85 maps and construct shapefiles that represent the planned transit network as of 1985.

2.1.3 Travel time simulation

To fit our model, we require travel times between all locations on all modes and all combinations of modes in each year. This is a large problem, complicated by the fact that we do not observe historical transit service and non-rapid transit is only partially observed. We first "backcast" modern infrastructure to prior periods. To do this, we begin with an OpenStreetMap (OSM) representation of the street network and remove highway additions. For rapid transit, we use a General Transit Feed Specification (GTFS) from 2017 and remove lines added since 2007 and 1994. For non-rapid transit, we combine modern GTFS (for formal buses) with two crowd-sourced databases of informal transit and simulate historical GTFS that include formal and informal transit. We then simulate travel times between all pairs of origins and destinations on all combinations of modes using the Rapid Realistic Routing on Real-world and Reimagined (R5) networks algorithm (Conway, Byrd, and Van Eggermond 2018). This method multi-modal travel and accounts for expected waits due to line or mode transfers. See the appendix for more details.

2.2 Summary Statistics and Context

Table 1 shows summary statistics using harmonized EOD data for each survey wave, restricted to tracts that we observe in all waves. Panel A contains modal split, average distance, and average time per trip conditional on mode. Because people can use multiple modes across or within trips, the modal split shares do not add up to one.[17] Panel B gives the share of trips that use only one, only two, or three or more modes (excluding walking and bicycling). Panel C reports vehicle ownership and usage statistics at the household level.

17. Moreover, the average times that condition on one mode are longer than those that do not because they draw from a subsample that includes more multimodal trips.

[Table 1 about here.]

Several patterns emerge from Table 1. The overall transit mode share has declined from 75% of trips in 1994 to 64% of trips in 2017. This is predominantly due to a decline in bus and *colectivo* use; the modal split of the subway has increased slightly, and BRT now has a sizable 6% share. Multimodality is important, comprising roughly 20% of trips in all years. Private vehicle and taxi use has also increased. Automobile adoption has increased over time: in 1994, 32% of households use a vehicle, while 39% used a vehicle in 2017.[18] Travel times have increased substantially, despite trip distances remaining relatively unchanged. This suggests substantial increases in congestion.

3 Empirical Design

We first describe a simple empirical model that is often used to study the effects of transit access on automobile adoption (e.g., Baum-Snow and Kahn 2000). We then extend the model to account for scale effects of transit, i.e., transit density. We then describe several supporting models that probe robustness to our identifying assumptions and to heterogeneous treatment effects.

Our baseline model for studying the effects of proximity on rapid transit is a staggered difference-in-differences specification. Let $\mathcal{T}_{\omega t}$ be the set of modes that household ω, which is observed in year t, reports using on their survey reference day. Households have adopted automobiles, i.e., $1[Veh. \in \mathcal{T}_{\omega t}] = 1$, if anyone in the household uses a vehicle.[19] Let D_{it} measure the proximity of census tract i to rapid transit in year t. The residential tract of household ω is tract $i = i_\omega$. Our main specification is given by the following equation:

$$1[Veh. \in \mathcal{T}_{\omega t}] = \beta D_{it} + \delta_i + \varsigma_t + e_{\omega it}. \tag{1}$$

Here, β captures the effect of proximity to rapid transit on household vehicle use. Residential tract fixed effects capture time-invariant characteristics of residential locations that may influence vehicle use (such as distance to the central business district or elevation). Residential tract fixed effects also capture average tract income from 1994 to 2017, controlling for persistent neighborhood segregation by income class.[20] Year fixed effects absorb factors that impact the city as a whole, such as average congestion levels, national wealth shocks, or automobile prices. We discuss identification assumptions for this baseline model in Section 3.1, as well as alternative research designs. Standard errors are clustered in two ways, by tract and by household. The additional clustering by household reflects the areal splitting of households during geonormalization.

18. Indeed, Guerra (2015) documents that between 1980 and 2010 the number of cars and light-duty trucks in Mexico City increased from 124 to 267 per thousand residents.

19. See footnote 16 for the robustness of this measure as a proxy for vehicle ownership.

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

We next turn to the specification that allows us to study scale effects of transit. We define scale of transit access in terms of the number of *lines* that are in proximity of a household's residential area, i. Lines typically serve different destinations, making them a reasonable measure of access.[21] In contrast, proximity to several stations on the same rapid transit line does not increase access to new locations.[22] Let $D_{it,(k)}$ be the normalized proximity of each line k of mode j, where (k) denotes a reversed order statistic (such that (1) is the most proximate of the indexed set). We modify Equation 1 to capture the effect the (k)th closest line as:

$$1[Veh. \in \mathcal{T}_{\omega t}] = \sum_{k=1}^{\bar{k}} \beta_{(k)} D_{it,(k)} + \delta_i + \varsigma_t + e_{\omega it}, \tag{2}$$

up to the $(\bar{k})$th closest line. Therefore, the effect of receiving a station that serves K lines in a residential tract relative to no rapid transit access is $\sum_{k=1}^{K} \beta_{(k)}$. We set $\bar{k} = 3$ given the distribution of lines in Mexico City, but also extend this up to $\bar{k} = 5$ in some specifications.[23]

We operationalize the notion of proximity to a rapid transit line with a piecewise functional form that is linear in negative normalized distance up to a maximum level, at which proximity becomes zero. Specifically, location i's proximity to rapid transit in year t is:

$$D_{it} \equiv \max \left\{ \frac{\bar{d} - \min_{s \in \mathcal{S}_t} d(i,s)}{\bar{d}}, 0 \right\} \in [0,1], \tag{3}$$

where d is the distance operator, $\bar{d}$ is a maximum distance for which an effect occurs, and $\mathcal{S}_t$ is the set of rapid transit stations operating in year t. For tracts that contain a rapid transit station in year t, $D_{it} = 1$, which then decreases to $D_{it} = 0$ for tracts more distant that $\bar{d}$.

To set $\bar{d}$, we estimate variants of Equation 1 with rapid transit use and vehicle use as outcomes, using 250m distance bins. Results (shown in Appendix Figure A3 and discussed in Appendix E.1) suggest a maximum effective radius of proximity of roughly 2km. If the magnitude of treatment effects decline in distance, setting $\bar{d}$ too high is conservative and attenuates estimates, while setting it too low inflates estimates (Butts 2023). This functional form provides the parsimony of a single treatment variable (per line), while also accommodating treatment effect heterogeneity by distance.[24]

21. Some lines run parallel to existing lines along some segments of their route. These lines may add redundancy to an existing network. We explore the effects of these lines on automobile adoption in Section 4.1.2.

22. Unlike other subway systems (such as in New York City), there are few instances of lines sharing tracks and tunnels in Mexico City's system.

23. Table A2 and A3 report the distribution and changes of tract-level transit line access in 1994 and in 2017 in our sample. Nearly 56% of tracts do not have or receive rapid transit by 2017. About 16% transition from no rapid transit access to one or more lines, with 14% receiving one line. About 14% of tracts have access to one rapid transit line in 1994; more than 8% increased to two or more lines by 2017. About 4% of tracts had access to three or more rapid transit lines in 1994; this increased to more than 14% by 2017. About 64% of tracts see no change in lines, 23% an increase of one line, 10% an increase of two lines, 3% an increase of three lines, and 0.5% an increase of four or more lines.

24. Appendix E demonstrates that using a binary measure of treatment attenuates estimates.

3.1 Identification and Supporting Designs

Identification of the baseline design (Equation 1) requires parallel counterfactual trends between tracts that receive a rapid transit line and those that do not. Put differently, the identifying assumption is that treated locations would have experienced similar average changes in outcomes in the absence of treatment as control locations experienced. This would potentially be violated if, for example, planners intentionally placed transit lines and stations in locations where they correctly anticipated transit demand would rise. Although we cannot directly test the parallel trends assumption given our data—we do not have enough early rounds of the EOD survey—we explore alternative research designs that are robust to specific ways in which this assumption could be violated. The first two alternatives reflect different assumptions about endogenous placement of transit stations. The last limits treatment-control comparisons to locations with similar ex ante levels of transit access, permitting more careful analysis of scale effects. In addition, we use auxiliary data to test for pre-trends in sociodemographic characteristics, as well as for sorting in response to transit expansions.

3.1.1 Supporting Design I: Arbitrary Municipality Effects

Suppose that the parallel trends assumption were violated because either different areas of Mexico City were on differing growth paths and transit expansions were concentrated in areas with higher or lower growth. Similarly, suppose that planners targeted transit expansions to serve particular regions. The first supporting design controls for these violations by replacing the year fixed effects in Equations (1) and (2) with municipality-by-year fixed effects (Muni-by-Year FEs):

$$1[Veh. \in \mathcal{T}_{\omega t}] = \sum_{k=1}^{\bar{k}} \beta_{(k)} D_{it,(k)} + \delta_i + \varrho_{mt} + e_{\omega it}. \tag{4}$$

In Equation 4, m refers to the municipality within which tract i lies. Our sample spans 31 municipalities (comprised of 16 alcaldías in Ciudad de México and 15 municipios in Estado de México). The inclusion of Muni-by-Year FEs limits the treatment and control comparisons to locations within the same municipality and is therefore robust to differential regional trends in economic behavior, such as broad growth in the southeastern and western portions of Mexico City.

3.1.2 Supporting Design II: PM85

For the next supporting design, we select control locations that could have plausibly received treatment but did not, as in Pollmann (2023). We use the *Plan Maestro 1985*, a planning document developed by the Mexico City government in the early 1980s in order to guide the continued development of the Metro. PM85 contains line and station locations for a network that would have covered much of Mexico City. Lines included in PM85 that were not built as planned are shown in

green in Figure 1.[25] This design takes as its sample a specific set of locations that were untreated in 1994 and were scheduled to receive a station according to PM85. Identification in this design assumes that locations that were intended to receive a transit line in PM85 but that did not actually receive transit have the same counterfactual trends as locations that received transit. Thus, this design assumes that treatment and control areas share a common trend if planners included them in PM85, even if they are in distant locations, while the Muni-by-Year design assumes that areas share a common trend if they are in close proximity.[26]

3.1.3 Supporting Design III: Scaffolded Design

Equation (2) is more demanding of the data than our baseline specification, as it differentiates the scale effects (or "dose") that each location received. This limits our ability to use the PM85 design to evaluate counterfactual changes to transit scale beyond access to the first line. However, we develop a third supporting design that limits comparisons by baseline level of treatment to estimate returns to scale.

We turn to a piecewise (or scaffolded) design that estimates the returns to different levels of treatment in the spirit of Callaway, Goodman-Bacon, and Sant'Anna (2024). This design limits the sample for estimating the effect of each additional line based on the number of lines present in the pre-treatment period. For example, to estimate the effect of access to a second rapid transit line, sample includes only the set of locations that had exactly one rapid transit line present. This model eliminates comparisons of newly treated units to more intensely treated locations and, like the PM85 design, to locations that have already been treated.

To be precise, first denote a scaffold as the set $\mathcal{S}(k') \equiv \{i : D_{i,t-1,(k')} = 0 \cup D_{i,t-1,(k'-1)} > 0\}$, with k' a positive integer. That is, the scaffold $\mathcal{S}(2)$ consists of all locations i that had no access to two transit lines last period (so that $D_{i,t-1,(2)} = 0$) and had some access to one transit line last period (so that $D_{i,t-1,(1)} > 0$). Indexing scaffolds by s, each scaffold can be estimated by the following model:

$$1[Veh. \in \mathcal{T}_{\omega t}] = \sum_{k=k'}^{\bar{k}} \beta_{(k)} D_{it,(k)} + \delta_{is} + \varsigma_{st} + e_{\omega ist}, \quad \forall i \in \mathcal{S}(k') \tag{5}$$

To facilitate hypothesis testing, we stack scaffolds into a single equation, clustering standard errors at the residential tract level (i.e., across scaffolds).

25. Zárate (2023) uses one of the proposed lines in the PM85 to test a model of how informality responds to transit. Figure A1 shows the map that we use to create the PM85 design.

26. This design excludes always-treated locations, which provides one way to mitigate concerns about negative weights for the TWFE estimates of Equation (1) (e.g., De Chaisemartin and d'Haultfœuille 2023).

3.1.4 Testing for Pre-Trends (in Auxiliary Data) and Sorting

Transportation infrastructure can induce sorting (Barwick et al. 2021). Sorting would pose issues to the interpretation of our results as a net increase in transit use. In the presence of sorting, our results could instead be consistent with transit users re-locating to transit-accessible areas without necessarily changing the city-wide level of transit use.

We check whether sorting on observables is an important issue in this setting. Table A11 shows the impacts of access to any rapid transit on households' observable characteristics. While sorting based on unobservable characterstics may still be present, absence of changes in observable characteristics in response to transit would suggest that sorting is not a first order concern. Indeed, the results in Table A11 Panel A show insignificant changes in observables, except in the case of household head age. People are similar across locations near rapid transit stations in terms of income, employment, gender, student status, and education level (High School and College). However, on average people near rapid transit stations are about one year older. Panels B and C report results using the PM85 and Muni-by-Year FE designs, respectively. These designs (and especially the Muni-by-Year FE design) show little evidence to support sorting on observables.[27]

4 Results

Table 2 reports estimates of baseline and scale effects of transit proximity on automobile adoption.[28] Column 1 reports the TWFE estimate of Equation (1) and suggests that proximity to one or more transit lines corresponds significantly with a 3.8 percentage point (pp) decline in household automobile adoption. These estimates are sizable compared to those of related studies. Baum-Snow and Kahn (2000) reports that decreasing residential distance to rail transit by 2km is associated with a 1.4pp increase in person-level transit use for commuting. Severen (2023) finds that residential proximity to rail transit by 1km in Los Angeles increases the commute share of rail by 0.8pp. Gaduh, Gračner, and Rothenberg (2022) find no effect of BRT with 1km on auto use.[29]

[Table 2 about here.]

Column 2 of Table 2 reports results for extending the baseline specification to account for scale

27. Zárate (2023) finds little evidence of changes in household composition after Line B is constructed, and more generally Gonzalez-Navarro and Quintana-Domeque (2016) find little evidence of changing household composition following road construction in Mexican cities. In general, there is less residential segregation in Mexican cities than in US cities (Monkkonen, Giottonini, and Comandon 2021). Table A13 shows income transitions in neighborhood rank in our data. About 39% of tracts are in the same quintile of average income in 2017 that they were in 1994, and 75% are in the same or adjacent quintile.

28. Throughout our results, we use an indicator for whether or not any household member used a vehicle as our preferred measure of automobile adoption. Other measures, such as vehicle ownership, are not available across all waves of our data, although we provide results on these other measures in the Appendix.

29. We explore heterogeneity between subway and BRT briefly in Appendix E.2 and more fully in our companion piece (Hernández-Cortés, Oliva, and Severen 2024).

effects (Equation 2). These results suggest a 3.0pp decline in automobile adoption from the first transit line, a larger marginal decrease of 10.8pp from the second transit line, and a yet larger marginal decrease of 11.4pp from a third transit line. These effects are marginal; adding them together yields a cumulative decline of 25.3pp in automobile adoption in response to constructing three transit lines near a residence (relative to zero transit lines). These results suggest substantial scale effects of transit on automobile adoption.

We provide tests of whether these scale effects represent constant returns to scale (CRS). Specifically, we define a null hypothesis of CRS for access to two lines (relative to one line) as $\beta_{(1)} + \beta_{(2)} = 2\beta_{(1)}$ and for three lines (relative to two lines) as $\beta_{(1)} + \beta_{(2)} + \beta_{(3)} = \frac{3}{2}\left(\beta_{(1)} + \beta_{(2)}\right)$. The bracketed numbers in Column 2 of Table 2 report the p-values of tests of these null hypotheses, and suggest increasing returns to scale for a second rapid transit line and CRS for a third line.

Columns 3 and 4 of Table 2 implement the Muni-by-Year design, accommodating arbitrary differences in regional trends. This more demanding specification soaks up some of the power of the baseline design, and the effect of a single transit line is no longer significant. However, the marginal effect of a second line is a significant reduction on automobile adoption of 5.2pp. The third line effect is no longer significant, but the cumulative effect of three or more lines is an 11.0pp decrease in automobile adoption. This design fails to reject the null of CRS in the density of transit lines.

Table 4 reports results from the other two designs. Column 1 shows the effects of receiving one or more lines transit lines relative to locations that were likely to have received a transit line but did not. Here, the average effect is small and insignificant. Column 2 reports results from the scaffolded design. These results are quantitatively in line with those from the basline design in Column 2 of Table 2 and exhibit substantial scale effects (though the evidence for increasing returns to scale is not present). The cumulative effect of three or more lines using the scaffolded design is a 24.6pp reduction in automobile adoption. Finally, Column 3 combines the Muni-by-Yr design with scaffolded design, including municipality-by-year-by-scaffold fixed effects.

[Table 3 about here.]

[Table 4 about here.]

These results collectively suggest at most small effects of a single rapid transit line on vehicle adoption. However, they consistently show a substantial marginal effect of a second transit line. Moreover, estimates of the cumulative effects of three of more rapid transit lines on vehicle adoption are all highly significant and vary between -11.0pp and -25.3pp. Taken together, this indicates that simply providing transit access is insufficient to appreciably shift household vehicle adoption. Rather, access to two or more rapid transit lines is required to change behavior.

4.1 Additional Treatment Specifications, Outcomes, and Robustness

The effects of rapid transit proximity could vary by mode (subway or BRT). While our companion paper (Hernández-Cortés, Oliva, and Severen 2024) comprehensively compares subway and BRT, differences in effects across mode could substantial alter interpreting the results in Tables 2 and 4 as scale effects, particularly is the mode of the first or second line were consistently subway or consistently BRT. We develop a test for differential scale effects of subway and BRT in Appendix E. Results (shown in Figure A4) indiciate that subway and BRT have strikingly similar returns to scale. Appendix E also compares our piecewise continuous measures of treatment, $D_{it,(k)}$, to discrete indicators, revealing that discrete measures attenuate estimates (as suggested in Butts 2023).

We also examine whether our results hold for vehicle ownership, rather than automobile adoption—the proxy for vehicle ownership we construct based on vehicle use at the household level. A disadvantage of this outcome is that vehicle ownership was not recorded in 1994, so estimates rely on only two cross sections instead of three. Results (in Appendix Table A4 and Table A5) are qualitatively similar to the main outcomes, although provide slightly different nuance and are based on different variation. In most designs, there is small and marginally significant effect on vehicle ownership for the first line, an insignificant effect for a second line, and sometimes a quite substantial effect for a third line. Results continue to suggest substantial scale effects of transit.

4.1.1 Robustness to Heterogeneous Treatment Effects

We provide evidence suggesting that our results represent a reasonable average of heterogeneous treatment effects. First, we note that the PM85 and scaffolded designs omit always treated units. That results from these design are broadly similar to the other designs mitigates concerns that negative weights confound our other estimates (De Chaisemartin and d'Haultfœuille 2023). Second, we estimate our main effects using long differences (i.e., omitting the 2007 cross section).[30] Long-difference results, shown in Appendix Table A6, are very similar to our headline results.

4.1.2 Heterogeneity across line functionality

In this section we explore whether the effects of proximity to multiple lines vary with the different types of functionality a line can provide. For example, tracts that gain proximity to line segments that run perpendicular to existing lines are more likely to gain *connectivity* to new locations. Peripheral tracts that gain proximity to rapid transit lines for the first time can gain *access* to denser areas of the city. Finally tracts that gain proximity to lines (or line segments) that run parallel to

30. We could also estimate all combinations of years, as in Goodman-Bacon (2021). Results (not reported) are relatively stable across years.

existing lines can expand capacity in a given route or provide *redundancy* in case of unexpected shocks. We explore heterogeneity along these dimensions to better understand which of these transit attributes increases the substitutability between transit and vehicles. Note that *connectivity* and *access* would be more important for expanding market access for consumption and work in the standard sense—by reducing average travel times to a wider set of valuable destinations—whereas *redundancy* could instead hedge households against sporadic shocks along well connected routes.

We first classify treatment tracts (those that are within line proximity) within these three categories $\mathcal{E} = \{\varnothing, \text{Access}, \text{Connectivity}, \text{Redundancy}\}$. We then estimate the following equation

$$1[Veh. \in \mathcal{T}_{\omega t}] = \sum_{j \in \mathcal{E}} \beta_\epsilon I_{j,it} + \delta_i + \varsigma_t + e_{\omega it}, \tag{6}$$

where $I_{j,ct}$ are indicators for each treatment type and $I_{\varnothing,ct} = 1$ represents a tract that is near transit but is not assigned to a particular treatment type. Some tracts experience more than one treatment at a time.

Results are shown in Appendix Table A8, and reveal that connectivity effects have, by far, the large impacts on automobile adoption. Access effects are not very large and not significant, consistent with the relatively small effects of the first line in Table 2. There is also little consistent evidence that redundant treatment drives automobile adoption.

5 Model: Non-Homothetic Demand for Mobility and Automobiles

Our results show that rapid transit expansions can induce substitution away from vehicles and towards transit. Importantly, we show that the density of the transit network is key to generating those substitution patterns: it takes residential access to more than one line to induce people to use public transit as opposed to private vehicles. Table A8 shows that these scale-of-network effects are likely driven by increased connectivity granted by additional lines, as opposed to redundant connections. These results suggest that further expansions of the network that increase connectivity will result in additional public transit use.

To better understand what is it about public transit expansions that reduces the relative value of vehicle ownership, we take advantage of the richness of our data to estimate a model where individuals have heterogeneous preferences for consumption and work destinations and for trip characteristics, such as time traveled and mode. This model allows us to decompose the reduced form vehicle use response that accompanies transit expansion into its effects on consumption and work market access, and it allows us to explore the role of preference heterogeneity in shaping these responses.

5.1 Model Setup

We are interested in vehicle ownership decisions, which are generally made at the household level. We assume individuals within a household share the same preferences (a unitary household model) and car ownership status. Thus, we use ω to index both households and household members. Household members consider the value of trips to work and consumption destinations when making decisions about car ownership. We assume they make this decision conditional on a level of potential income, y_ω, a working status, $l_\omega \in \{0, 1\}$, and a residential location, i. We also assume individuals have an initial work location, j, when making their car ownership decision. However, they also consider the location of alternative jobs in the event that they face a job transition shocks. A household is a working household, $l_\omega = 1$, if at least one member of the household is part of the labor force.[31]Non-working households do not consider work locations in their car ownership decision, and only value consumption destinations.

Our model assumes that household preferences over workplace, consumption destinations and mode choice vary by a life-cycle adjusted measure of permanent income, which we call "potential income". Our implementation of potential income, described in Section 6.1, limits endogeneity of income with respect to workplace choice Zárate (2023) and overcomes missing income data in 2017. Potential income for the household, y_ω, corresponds to that of the member within the household with the highest potential income (see Section 5.2).[32] We adopt a variant of the latent class structure for modeling heterogeneity in preferences. Like the typical latent class model, we assume households belong to one of a discrete number of classes: $g \in \{1, 2, 3\}$. However, unlike the latent class model, class assignment probabilities can depend on observable characteristics of households (Bhat 1997).

Specifically, we use an ordered logit process to flexibly assign households to groups based on their potential income, y_ω. The probabilities of a household belonging to each of the three groups are given by

$$\begin{aligned}
\Pr(\omega \in g = 1 | i, y_\omega) &= \frac{1}{1 + e^{-(\mu_1 - \tilde{y}^*_\omega)}} \\
\Pr(\omega \in g = 2 | i, y_\omega) &= \frac{1}{1 + e^{-(\mu_2 - \tilde{y}^*_\omega)}} - \frac{1}{1 + e^{-(\mu_1 - \tilde{y}^*_\omega)}} \\
\Pr(\omega \in g = 3 | i, y_\omega) &= 1 - \frac{1}{1 + e^{-(\mu_2 - \tilde{y}^*_\omega)}},
\end{aligned}$$

31. The EOD does not distinguish between being unemployed and being out of the labor force. However, unemployment rate in Mexico is low during our period of study, so we assume households where no members are employed are out of the labor force or non-working.

32. For working households, household potential income is assigned based on the *working member* of the household with the highest potential income. For non-working households we simply assign the highest potential income within the households to all household members.

where y^*_ω, and thus all three probabilities, vary continuously with potential income.[33] This latent class structure allows collapsing continuous preference heterogeneity by income into heterogeneity by group, g.

In our model, a representative household member sequentially chooses vehicle ownership, workplace or consumption locations, and a combination of modes of travel to each of these locations.[34] At each decision point, household members are uncertain of what the following decision will be, and thus choose based on the expected value of the subsequent choice set. Households choose vehicle ownership $a \in \{0, 1\}$ based on access to their current workplace, expected market access to future workplaces, and expected market access to consumption destinations. Conditional on vehicle ownership, household members independently choose destinations $j \in \mathcal{J}$ based on the value of working or consuming in that location net of expected commuting costs. Conditional on vehicle ownership, trip type $c \in \{\text{Work, Cons.}\}$, and workplace or consumption destination, households choose a combination of modes $k \in \mathcal{K}_a$ for each trip. Households may have different preferences over travel time and modes for work and consumption trips. The choice set $\mathcal{K}_a$ consists of ordered combinations of modes, allowing for multimodality, and is indexed by a to represent differing choice sets based on household vehicle ownership.

Potential income flexibly enters the vehicle choice decision and also determines latent class assignment. However, for the workplace and mode choice problems, group membership probabilities are a sufficient statistic for income.

We consider each of these choices in turn, starting at the 'bottom' of the problem (mode choice) and building back up to vehicle choice. We present the general case in each section below, and note alternative and simpler assumptions as well.

5.2 Mode Choice

In a slight abuse of notation, we use ω to index, both, trips and representative household members when discussing mode choices and destination choices. In the data, we often observe multiple trips per household member and/or household. Conditional on potential income y_ω, these trip observations are treated as independent.

For each trip, household members choose a sequence of modes conditional on destination, trip type, and vehicle ownership. Each sequence of modes, k, is selected from choice set $\mathcal{K}_a$, where $\mathcal{K}_1 = \{\mathcal{K}_0 \cup \text{car}\}$. The non-car choice set $\mathcal{K}_0$ contains ordered combinations of up to two modes

33. We define y^*_ω as:

$$y^*_\omega = \tilde{y}^*_\omega + \vartheta_\omega = Q(F(y_\omega); \mu_f, \sigma_f) + \mu_{mi} 1_{mi} + \vartheta_\omega,$$

where ϑ is distributed iid logistic, F is the empirical cdf of the potential income distribution, and Q is the log-normal quantile function with mean and variance parameters μ_f and σ_f. We include the parameter μ_{mi} to assign income for households that report working (and for which we see a commute) but who do not report their income. We allow these missing income terms to vary by year.

34. Recall that household members share the same preferences and car ownership within the household.

from $K \equiv \{\text{subway, BRT, informal}\}$ along with taxi, so $\mathcal{K}_0 \equiv \{K^2 \cup \text{taxi}\}$.[35] We refer to the set K^2 as public transit.

Thus, for each trip, a representative household member choosing set $\hbar$ receives indirect (random) utility

$$\mathcal{U}^c_{\hbar jg,\omega}(\tau_{ij\hbar}) = b^c_{\hbar g} + \sum_{k \in \{\text{transit,taxi,car}\}} 1_k \kappa^c_{gk} \tau_{ij\hbar} + \nu^c_{ijg\hbar,\omega},$$

where $b^c_{\hbar g}$ are utility shifters that capture the group-specific mean utility (net of fixed costs) associated with each mode sequence, $\hbar$, and trip type, c. Uncongested travel time is $\tau_{ij\hbar}$, and we allow for a different marginal disutility of travel time $\kappa^c_{gk} < 0$ for transit, car, and taxi.[36] We also allow the marginal disutility of time to differ for work and consumption trips. Finally $\nu^c_{ijg\hbar,\omega}$ captures idiosyncratic preferences for mode sequences, and unobservable trip shocks – such as unexpected delays. We allow for correlation in $\nu^c_{ijg\hbar,\omega}$ across modes. Specifically, we assume a nested logit structure with several layers of nests. The first nest corresponds to the choice between public transit and private modes (car and taxi). Within the public transit nest, choices are correlated with parameter λ_{Public} and agents decide if their first trip will be on rapid transit (subway and BRT) or informal transit. Within the rapid transit nest, choices are correlated with parameter $\lambda_{RapidTransit}$ and agents decide whether they will first take subway or BRT. Finally, conditional on their first transit choice, choices are correlated with parameter $\lambda_{LastMode}$ and agents decide if they will take a second mode and if so, which mode. Within the private transit nest, choices are correlated with parameter $\lambda_{Private}$ and agents choose between taking their household car (if they have one) and a taxi. We graphically present the nesting structure in Figure 2.

[Figure 2 about here.]

5.3 Workplace and Consumption Location Choices

Households choose workplace and consumption destinations based on expected travel costs, $\ln(d^c_{iajg}) \equiv \mathbb{E}\left[\max_{\hbar \in \mathcal{K}_a | iaj} \mathcal{U}^c_{\hbar jg,\omega}(\cdot)\right]$, and the value of the destination, ϖ^c_{jg}, where c denotes destination purpose (Work or Consumption): $c \in \{W, C\}$. As in Tsivanidis (2023), d^c_{iajg} can be interpreted as the expected optimized travel time between two locations. This interpretation acknowledges that chosen mode can change depending on the realization of unobserved shocks, $\nu^c_{ijg\hbar,\omega}$.

We model costs and benefits associated with each of the two trip purposes differently. We assume that working households are endowed with a workplace when we observe them in each cross-section of the survey. However, their decision to own a car also depends on the possibility

35. Only a very small share of trips use more than two modes.

36. The different travel-time disutilities across modes allows us to capture differences in variable costs (e.g, taxi rate per km, gasoline prices). In the estimation, we let these vary year by year to changes in congestion and time-varying variable-use costs.

of changing workplaces in the future, which happens with probability p.[37] This dynamic structure allows for flexibility in the order of car ownership and workplace decisions. If workplace transitions occur often, car ownership should hinge on the portfolio of workplaces that are accessible through driving. But if workplace transitions are rare, car ownership should largely ignore other workplace possibilities and hinge on current workplace location.[38] In contrast with the work purpose location decisions, we assume consumption location decisions are always made after the car ownership decision is made. Thus, when considering the car purchase, households value the portfolio of choices and are not committed to specific consumption destinations. This asymmetry will be reflected in the form of the sub-utilities associated with trips for work and consumption purposes that enter the car ownership choice.

Specifically, we define the indirect random sub-utility associated with work commutes given residential location i, car ownership a, latent class g, and workplace location j as

$$\begin{aligned}\mathcal{V}^W_{iajg,\omega}\left(\varpi^W_{jg}, d^W_{iajg}\right) =& \frac{1}{\epsilon^W_g}\left(\ln \varpi^W_{jg} + \epsilon^W_g \ln d^W_{iajg} + e^W_{ijg,\omega}\right) \\ &+ r\left(p\mathbb{E}\left[\max_k \mathcal{V}^W_{iakg,\omega}\left(\varpi^W_{kg}, d^W_{iakg}\right)\right] + (1-p)\,\mathcal{V}^W_{iajg,\omega}\left(\varpi^W_{jg}, d^W_{iajg}\right)\right),\end{aligned} \tag{7}$$

where r is the discount factor and an idiosyncratic component of indirect utility, $e^W_{ijg,\omega}$, is distributed Extreme Value Type I.[39] The second line in (7) corresponds to the continuation value which is realized after getting a new draw of the workplace-transition shock. With probability p, households have to re-optimize across workplace locations, conditional on place of residence, i and car ownership a. However, with probability $(1-p)$ their workplace location (and job) is the same they are endowed with at the moment we observe them in our sample. Thus, Equation 7 is defined recursively and—using the EV Type I distributional assumption—can be solved for $\mathcal{V}^W_{iajg,\omega}\left(\varpi^W_{jg}, d^W_{iajg}\right)$:

$$\mathcal{V}^W_{iajg,\omega}\left(\varpi^W_{jg}, d^W_{iajg}\right) = \frac{r^*}{\epsilon^W_g}\ln \varpi^W_{jg} + r^* \ln d^W_{iajg} + \frac{r^*}{\epsilon^W_g}\frac{rp}{(1-r)}\ln\left(\Phi^W_{iag}\right) + \tilde{e}^W_{ijg,\omega}, \tag{8}$$

where $r^* = \frac{1}{1-r(1-p)}$ and $\tilde{e}^W_{ijg,\omega} = \frac{r^*}{\epsilon^W_g} e^W_{ijg,\omega}$.[40] The term Φ^W_{iag}, where $\Phi^W_{iag} = \sum_k \varpi^W_{kg}\left[d^W_{iakg}\right]^{\epsilon^W_g}$, has the interpretation of commuting (labor) market access; i.e., it reflects the value of the portfolio of workplace choices that are accessible from location i with car ownership a for latent class g (e.g.,

37. Using quarterly social security data from Mexico, Samaniego de la Parra and Fernández Bujanda 2024 find that 81.7% of all formal employees stay at the same establishment with the same formality status. At the same time, the probability of job separation is approximately 7.5-7.7% for workers.

38. Non-working households (e.g., retired) do not have an existing workplace and do not choose workplaces at any point.

39. The idiosyncratic term $e^W_{ijg,\omega}$ can capture any household-job specific shifters, such as match-specific productivity, preferences for job attributes, etc.

40. See Appendix D.1 for derivations.

Ahlfeldt et al. 2015).

The random sub-utility for consumption trips is given by

$$\mathcal{V}^C_{iajg,\omega}\left(\varpi^C_{jg}, d^C_{iajg}\right) = \frac{1}{\epsilon^C_g}\left(\ln \varpi^C_{jg} + \epsilon^C_g \ln d^C_{iajg} + e^C_{ijg,\omega}\right). \tag{9}$$

As in Equation 7, $e^C_{ijg,\omega}$ is assumed to be distributed Extreme Value Type I. However, this term also captures trip-specific shocks to destinations that are drawn anew before every trip. As such, it represents variation in the purpose of consumption trips that could lead to difference destinations (e.g., grocery stores versus dining).

The distributional assumptions allow deriving quantities of households that choose each destination. The share of group g who chooses destination j for trip type c, conditional on car ownership a and residential location i, is

$$\pi^c_{j|iag} = \frac{\varpi^c_{jg}\left[d^c_{iajg}\right]^{\epsilon^c_g}}{\sum_k \varpi^c_{kg}\left[d^c_{iakg}\right]^{\epsilon^c_g}}. \tag{10}$$

That is, choice probabilities match those given by a model of destination choice with idiosyncratic preferences over locations distribution Fréchet with shape parameter ϵ^c_g, as is common in quantitative trade models (e.g., Eaton and Kortum 2002; Redding and Rossi-Hansberg 2017).[41]

5.4 Car Ownership Decision

Households choose car ownership at the top decision point of the model based on its cost and the sub-utilities associated with work and consumption trips for a representative member of the household. For work travel, the relevant sub-utility is given by Equation 8. Note that this sub-utility includes the household-specific work-related shock, $\tilde{e}^W_{ijg,\omega}$, because we assume car ownership decisions are conditional on the current workplace.

However, we assume household members are not committed to a single consumption destination. Instead, they maximize across locations every time they make a consumption trip after observing the realizations of shocks, $e^C_{ijg,\omega}$. Thus, when choosing whether or not to own a car, they consider the entire portfolio of destination choices that are accessible with and without it. Conditional on residential location i, car ownership a, and latent class g, the expected value of access to destinations for each trip type can thus be written succinctly as:

$$V^C_{iag} \equiv \ln \mathbb{E}[\max_j \mathcal{V}^C_{iajg,\omega}\left(\varpi_{jcg}, d_{iajcg}\right)] = \ln\left[\Phi^C_{iag}\right]^{1/\epsilon^C_g} = \ln\left(\sum_j \varpi^C_{jg}\left[d^C_{iajg}\right]^{\epsilon^C_g}\right)^{1/\epsilon^C_g} \tag{11}$$

41. Derivations for each $c \in \{W, C\}$ are in Appendix D.1.

due to the distributional assumption on $e^C_{jg,\omega}$. Here, Φ^C_{iag} reflects consumption market access, and V^C_{iag} is a function of $\left[\Phi^C_{iag}\right]^{1/\epsilon^C_g}$ is the group-specific value of access to all destinations for consumption. Thus, we can write the indirect utility for each car ownership choice, $a \in \{0,1\}$, as

$$\mathcal{W}_{iajg,\omega}\left(p_{\text{car}}a, \mathcal{V}^W_{iajg,\omega}, V^C_{iag}, l_\omega\right) = \alpha_g p_{\text{car}} a + \beta^W_g l_\omega \mathcal{V}^W_{iajg,\omega} + \beta^C_g V^C_{iag} + \bar{\eta}_{iag} + \eta_{iag,\omega}, \tag{12}$$

where α_g captures the utility-equivalent cost of car ownership, β^c_g is the relative value of market access for $c \in \{W, C\}$ type trips, $l_\omega \in \{0,1\}$ denotes working status of the household, $\bar{\eta}_{iag}$ captures common features of location i that may favor car ownership or transit for a particular group, and $\eta_{iag,\omega}$ are idiosyncratic preferences for car ownership. Notably, each preference term α, β^c, and $\bar{\eta}$ may vary with income through g.

Under a standard logit assumption on $\eta_{iag,\omega}$, solving the problem $\max_{a\in\{0,1\}} \mathcal{W}_{iajg,\omega}(\cdot)$ generates the following function of choice probabilities:

$$\ln\left(\frac{\Pr(a=1|i,j,g)}{\Pr(a=0|i,j,g)}\right) = \beta^W_g l_\omega \left(\mathcal{V}^W_{i1jg} - \mathcal{V}^W_{i0jg}\right) + \beta^C_g \left(V^C_{i1g} - V^C_{i0g}\right) + \tilde{\eta}_{ig}, \tag{13}$$

where $\tilde{\eta}_{ig} = \bar{\eta}_{i1g} - \bar{\eta}_{i0g} + \alpha_g p_{\text{car}}$ represents unobserved factors that may make vehicle ownership more or less valuable in tract i.

Substituting Equations (8) and (9) into Equation (13) and simplifying terms yields

$$\begin{aligned}\ln\left(\frac{\Pr(a=1|i,j,g)}{\Pr(a=0|i,j,g)}\right) =& \beta^W_g l_\omega r^* \ln\left(\frac{d^W_{i1jg}}{d^W_{i0jg}}\right) + \beta^W_g l_\omega \frac{r^*}{\epsilon^W_g}\frac{rp}{(1-r)} \ln\left(\frac{\Phi^W_{i1g}}{\Phi^W_{i0g}}\right) \\ &+ \beta^C_g \frac{1}{\epsilon^C_g} \ln\left(\frac{\Phi^C_{i1g}}{\Phi^C_{i0g}}\right) + \tilde{\eta}_{ig},\end{aligned} \tag{14}$$

where the terms $\ln \varpi^W_{jg}$ and $\tilde{e}^W_{ijg,\omega}$ cancel out when difference across car ownership choices.

6 Model Estimation and Results

We next describe how we implement the model, identification of model parameters, and model results. Notably, the model description in Section 5 is presented for a single time period. When implementing the model, we exploit the three cross-sections of travel survey data to aid identification. To this end, we index periods with t to reflect survey years. Where applicable, we also contrast the latent-class model with a simpler single-class version of the model that collapses class membership to a single type.

In principle, the full latent class model could be estimated in a single step via an expectation-maximization algorithm using the joint likelihood across all decisions to derive the objective func-

tion. Such an approach would utilize all observed behavior to assign households to latent classes. However, the large number of fixed effects in the workplace and consumption location problems, as well as in the car ownership problem, suggest that such an approach would be computationally intensive. We therefore use just the mode-choice model to assign latent classes, and then treat probabilistic class assignment as fixed for the workplace, consumption location, and car ownership problems.

6.1 Potential Income

We create a measure of life-cycle adjusted permanent income, which we call "potential income". Studying heterogeneity in response by income is a first order concern, as responses to transit vary strongly with income (Glaeser, Kahn, and Rappaport 2008; LeRoy and Sonstelie 1983). However, there are two challenges in using reported earned income to stratify heterogeneous response. First, income may respond endogeneously to the presence of transit as transit connects workers with new employment opportunities (Zárate 2023). A second challenge in our setting is that income data were not collected in the 2017 EOD.

We construct our proxy of life-cycle adjusted potential income using individual characteristics (gender, age, education) and residential location, but not workplace. We first estimate Mincerian models of wage for each EOD wave that condition on pre-determined worker characteristics and residential locations. We then use predicted wage, denoted y_ω, as a measure of potential (or life cycle-adjusted permanent) income. To overcome the absence of observed wage in the 2017 EOD, for this wave we predict income using individual data, harmonized characteristics, and residential location from 2015 *Conteo* confidential microdata and then predict wage using these coefficients and the 2017 EOD worker characteristics. See Appendix C for details, summary statistics, and validation.

6.2 Mode Choice Estimation and Results

The first estimation step parameterizes the mode choice and latent class models. This step recovers parameters of the trip mode choice sub-utility function, $\{b, \kappa, \lambda\}$, and latent class assignment parameters, $\{\mu, \sigma\}$. Estimation uses all trips that begin at home, segregated into work ($c = W$) and non-work ($c = C$) trips based on their stated purpose.[42] For τ, we use simulated uncongested travel times and add two parameters to accommodate cases when the travel engine failed to return travel times.[43] We normalize the level of utility to that of the subway-only choice.

We estimate the mode choice and latent class assignment models using a three stage procedure. The first stage separately estimates three single-class mode choice models on distinct subsets of the

42. The survey data are less geographically precise for trips that do not begin at home, so we focus on such trips. The majority of trips appear to be relatively symmetric.
43. See Appendix B for discussion.

population that fall within narrow income bands. The first band is between 1st and 11th centiles of potential income, the second between the 40th and 60th centiles, and the third band between the 89th and 99th centiles. In the second stage, we estimate the latent class assignment parameters holding fixed the parameters from the first stage. In the final stage, we jointly update the mode choice and latent class assignment parameters via penalized maximum likelihood (Firth 1993). The penalization uses estimates from the first two stages as initial values and for regularization. That is, instead of maximizing the log-likelihood $\mathcal{L}(\theta)$, where $\theta = \{b, \kappa, \lambda, \mu, \sigma\}$, we maximize the penalized function:

$$\mathcal{L}(\theta) - (\theta - \theta_0)R(\theta - \theta_0)',$$

where θ_0 are the initial parameter values estimated in the first two stages and R is the penalization matrix.[44] We set $R = \lambda I$ where I is conformable with θ and λ controls the strength of penalization.[45] Regularization is necessary to ensure interpretable results for each class, else estimates for one of the classes tend to diverge to extreme values.

Results from the full latent-class model are shown in columns 2a–2f of Appendix Table A15; columns 1a–1b present results from the corresponding single-class model.[46] Although this table contains many interesting findings, we briefly summarize the estimates and patterns that are most relevant to downstream model stages. Estimates of κ reveal that the marginal disutility of travel time is larger for work trips than for consumption trips, and generally is increasing in income. Moreover, preferences for transit modes are quite correlated, suggesting substantial substitutability.[47] This is particularly so for rapid transit modes. Notably, these parameters are relatively consistent across income groups. The mean utility parameters reveal that, ceteris paribus, BRT is slightly preferred to subway across all income groups. Interestingly, BRT is mildly normal relative to the subway, while informal modes are inferior relative to subway. All travelers, and in particular high-income travelers, receive substantial disutility from multimodal trips, reflecting a sizable transfer penalty.

The bottom panel of Appendix Table A15 reports latent-class assignment probabilities at different potential incomes. Households with the lowest potential incomes are assigned to the low-income class with a 91% probability, to the middle income class with a 9% probability, and a negligible probability to the high income class. Households at the median potential income are assigned to the middle latent class with a 79% probability, and to the low and high classes with 13% and 8% probabilities, respectively. The highest potential income households are 98% high-income class, 2% middle-income class, and negligibly low-income class. Together, these assignment prob-

44. This objective function with quadratic loss is numerically identical to Bayesian framework with a normal prior centered at θ_0 and common variance proportional to R.

45. We currently set $\lambda = 0.01$ via experimentation. Cross-validation is an alternative approach.

46. We estimate the single-class mode choice model in a straightforward manner via maximum likelihood.

47. Note that the correlation parameters λ are inversely related to correlation, so $\lambda = 1$ implies no correlation, $\lambda \in [0, 1)$ implies positive correlation, and $\lambda \to 0$ implies perfect correlation.

abilities reveal that the latent classes are indeed representing different segments of the population by income.

6.3 Workplace Choice Estimation and Results

We first discuss calculation of the market access terms (Φ^c_{iagt}) in Equations (9) and (11). We aggregate destination choices by latent classes by summing across latent class membership probabilities. The empirical share of each latent class g that chooses to travel to destination j conditional on car ownership a and residential location i is:

$$\pi^c_{j|iagt} = \frac{\sum_{\omega \in iact} \Pr(\omega \in g|i, y_\omega) 1[\omega \text{ chooses } j|iact]}{\sum_{\omega \in iact} \Pr(\omega \in g|i, y_\omega)}$$

where $\omega \in iact$ denotes households that have selected ia and are considering trip c in year t.

We then estimate the ϖ_{jcg} and ϵ_{cg} terms using a gravity framework, and then aggregate them to create market access terms. Specifically, we estimate the empirical analog of Equation (10), pooling data across years:

$$\ln(\mathbb{E}[\pi^c_{j|iagt}]) = \tilde{\omega}^c_{jgt} - \epsilon^c_g \ln d^c_{iajgt} - g^c_{iagt} \tag{15}$$

where t indexes years, $\tilde{\omega}^c_{jgt} = \ln(\varpi^c_{jgt})$ are workplace-by-class fixed effects, and $g^c_{iagt} = \ln(\Phi^c_{iagt})$ is the log denominator of Equation (10). In Equation (15), ϵ^c_g is identified both from dyadic variation in expected travel costs and changes in expected travel costs over time. Because of the presence of zero-valued flows, we estimate Equation (15) using Poisson pseudo-maximum likelihood.

Table 5 reports estimates of ϵ^c_g for both the single-class and latent-class models. This parameter is both a measure of the homogeneity of idiosyncratic preferences for workplace (where a larger value indicates more homogeneity and substitutability between places) and is also the extensive-margin elasticity of labor supply. The single-class estimate of this elasticity for workplace is 3.6, suggesting that workplaces are somewhat, but not completely, substitutable. The corresponding elasticity for consumption trips is estimated to be 5.0, suggesting that consumption destinations are more substitutable than workplaces. The workplace elasticities values are roughly in line with Severen (2023) and Tsivanidis (2023), but somewhat smaller than Ahlfeldt et al. (2015).[48] Moreover, these elasticities are quite comparable to Miyauchi, Nakajima, and Redding (2021), who use smartphone data and find workplace heterogeneity parameterized at 3.04 and itinerary heterogeneity of 4.5 (which is somewhat analogous to consumption heterogeneity in our framework).

[Table 5 about here.]

Our model also allows us to distinguish heterogeneity by latent class. Heterogeneity in prefer-

48. Ahlfeldt et al. (2015) use second moments on relatively few data points to estimate this elasticity.

ences for both workplaces and consumptions locations declines with income, and is substantially lower at the highest levels of income. The heterogeneity parameter across workplaces is 3.0 for the low-income class, rises to 3.6 for the middle-income class, and falls to 2.1 for the high-income class. These are not too different from Tsivanidis (2023), who finds an elasticity of 3.98 for low-skill workers and 2.65 for high-skill workers. There is notable heterogeneity in the idiosyncratic valuation of consumption destinations by income, however. The low-income estimate of 8.6 suggests that low-income households are highly homogenous in preferences for consumption destinations. High-income households, in contrast are much more hetegeneous, with a Fréchet value of 3.1. Middle-income households are in between, at 5.1.

6.4 Auto Demand Estimation and Results

Equation 14 quantifies the car ownership decision in terms of log-odds and returns values of β_g^c that can be interpreted as elasticities. However, this model uses parameters that are difficult to identify from cross-sectional data on vehicle use (like r and p) and is difficult to estimate with many fixed effects.[49] We therefore calibrate $r = 0.89$[50] and $p = 0.27$[51] to match annualized values from the literature (ECLAC 2015; Samaniego de la Parra and Fernández Bujanda 2024), and define

$$\Theta_{ijgt}^W = r^* \ln\left(\frac{d_{i1jgt}^W}{d_{i0jgt}^W}\right) + \frac{r^*}{\epsilon_g^W}\frac{rp}{(1-r)}\ln\left(\frac{\Phi_{i1gt}^W}{\Phi_{i0gt}^W}\right) \quad \text{and} \quad \Theta_{igt}^C = \frac{1}{\epsilon_g^C}\ln\left(\frac{\Phi_{i1gt}^C}{\Phi_{i0gt}^C}\right).$$

We then estimate the following linear-probability model:

$$1[Veh. \in \mathcal{T}_{\omega t}] = \sum_g \Pr(\omega \in g|i, y_\omega)\left(\mathcal{B}_g^W l_\omega \Theta_{ijgt}^W + \mathcal{B}_g^C \Theta_{igt}^C + \tilde{\eta}_{ig}\right) + \eta_{il_\omega} + \delta_t + e_{\omega it}. \tag{16}$$

Recall that the $\tilde{\eta}_{ig}$ capture time-invariant, class-specific residential factors that influence vehicle choice. We also include worker-status by residential location fixed effects η_{il_ω} to control for constant factors that may determine vehicle use and correlate with worker status (such as a large concentration of retirees) and δ_t to control time-varying common factors influencing auto adoption. Equation 16 recovers $\mathcal{B}_g^c$, which are semi-elasticities that map directly to β_g^c in Equation 14.[52]

Figure 3 shows estimates of combinations the semi-elasticities $\mathcal{B}_g^c$ across the income distribution; lines show point estimates and shaded areas show 95% confidence intervals. This figure reveals that work market access is relatively important for low-income households' automobile adoption decision. This importance decline slightly with income. In contrast, consumption access

49. See, e.g., the fractional logit correlated random effects estimator of Papke and Wooldridge (2008).

50. According to ECLAC (2015), the pre-2015 social discount rate in Mexico is 12%. Therefore, we use $r = \frac{1}{1+.12} = 0.89$.

51. According to Samaniego de la Parra and Fernández Bujanda (2024), the probability of job separation is approximately 7.5-7.7% for workers. We use 7.5% and calculate the annual separation rate as $p = 1 - (1 - 0.075)^4 = 0.27$

52. Specifically, $\beta_g^c = (\Pr(a = 1|i, j, g)(1 - \Pr(a = 1|i, j, g)))^{-1}\mathcal{B}_g^c$; see Appendix D.1.

is relatively unimportant for low-income households, but increases rapidly as income increases. For high-income households, consumption market access is substantially more important than work market access.

[Figure 3 about here.]

Also notable in Figure 3 is that neither effect is close to zero at the top of the income distribution. This suggests that relative market access would continue to play a key role in automobile adoption decisions even as incomes rise.

6.5 Understanding Scale Effects

To understand the origin of scale effects, we create structural analogs to the baseline research design presented in Equation 2. We first quantify how transit improvements impact relative market access terms. Specifically, we estimate the effect of changes in transit density separately for work and consumption relative market access for each class:

$$\left.\begin{matrix}\Theta^{W}_{ijgt} \\ \Theta^{C}_{igt}\end{matrix}\right\} = \sum_{k=1}^{\bar{k}} \gamma^{c}_{g,(k)} D_{it,(k)} + \delta_{ig} + \varsigma_{gt} + e_{\omega it}, \tag{17}$$

where the $\gamma^{c}_{g,(k)}$ capture the average effects of transit on changes in relative market access.

[Figure 4 about here.]

Figure 4 shows estimates of $\gamma^{c}_{g,(k)}$ across the income distribution, where the effect at each income y is $\gamma^{c}_{(k)}(y) = \sum_{g} \Pr(\omega \in g|y)\gamma^{c}_{g,(k)}$. Lines are darkest for the first line, and become lighter as transit density increases (up to $\bar{k} = 3$). These estimates reveal how access changes in response to transit, and reflect changes in travel time, mode availability, and the value of destinations. The effects of transit on relative work market access are markedly larger for higher-income workers, but do not increase with transit density. In contrast, there are mild improvements in work relative market access for low-income workers that increase slightly in transit density. Effects of transit on consumption relative market access are large and increase with scale across all income levels, but especially for lower-income workers

The effects depicted in Figures 3 and 4 reveal two different mechanisms driving the decision to purchase or not to purchase an automobile. They can be aggregated to derive a model-imputed total effect of transit density on vehicle ownership at each income level. Specifically, we calculate

$$\rho_{(k)}(y) = \sum_{g} \Pr(\omega \in g|y) \left(\mathcal{B}^{W}_{g} \gamma^{W}_{g,(k)} + \mathcal{B}^{C}_{g} \gamma^{C}_{g,(k)} \right).$$

We show these surfaces in Figure 5. These results show that the model successfully captures a substantial portion of the scale effects of transit. However, it is indicates that perhaps one line is

even less impactful than our results suggest. This could be the case if, for example, development in places with zero or one transit line is more auto-centric and less pedestrian and transit friendly than places with two or more lines.

[Figure 5 about here.]

We also estimate a single-class model, and present the results succinctly in Table 6. As in the case of the latent-class model, there are scale effect of transit on consumption access but not on workplace access (columns 1–2). The single-class model reports a higher semi-elasticity of work market access than consumption market access on vehicle adoption (column 3). In fact, comparing Table 6 to Figure 3, the single-class model reports semi-elasticities that correspond to those from the lower end of the income distribution from the latent-class model. This highlights the important of capturing heterogeneous preferences for understanding the role of transit on automobile adoption.

[Table 6 about here.]

7 Conclusion

We provide new evidence that transit has important scale effects when the determine its effects on automobile adoption. Using data from Mexico City that covers a period of extensive rapid transit expansion, we show that access to two or three rapid transit lines has a much more substantial effect on automobile adoption than access to a single line. We build a structural model to rationalize these findings and explore mechanisms. The model reveals consumption market access is the key driver of the scale effects of transit. Notably, these effects do not go to zero as incomes increase, suggesting a continued vital role for transit in urban policy.

References

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. 2015. "The economics of density: Evidence from the Berlin Wall." *Econometrica* 83 (6): 2127–2189.

Almagro, Milena, Felipe Barbieri, Juan Camilo Castillo, Nathaniel G Hickok, and Tobias Salz. 2024. *Optimal Urban Transportation Policy: Evidence from Chicago.* Technical report. National Bureau of Economic Research.

Balboni, Clare. 2019. "In harm's way? infrastructure investments and the persistence of coastal cities." PhD diss., London School of Economics and Political Science.

Balboni, Clare, Gharad Bryan, Melanie Morten, and Bilal Siddiqi. 2020. *Transportation, gentrification, and urban mobility: The inequality effects of place-based policies.* Technical report. National Bureau of Economic Research.

Barwick, Panle Jia, Shanjun Li, Andrew R Waxman, Jing Wu, and Tianli Xia. 2021. *Efficiency and equity impacts of urban transportation policies with equilibrium sorting.* Technical report. National Bureau of Economic Research.

Baum-Snow, Nathaniel. 2007. "Did highways cause suburbanization?" *Quarterly Journal of Economics* 122 (2): 775–805.

Baum-Snow, Nathaniel, and Matthew E Kahn. 2000. "The effects of new public projects to expand urban rail transit." *Journal of Public Economics* 77 (2): 241–263.

Baum-Snow, Nathaniel, Matthew E Kahn, and Richard Voith. 2005. "Effects of urban rail transit expansions: Evidence from sixteen cities, 1970-2000." *Brookings-Wharton Papers on Urban Affairs,* 147–206.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile prices in market equilibrium." *Econometrica* 63 (4): 841–890.

Bhat, Chandra R. 1997. "An endogenous segmentation mode choice model with an application to intercity travel." *Transportation science* 31 (1): 34–48.

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

Binelli, Chiara, and Orazio Attanasio. 2010. "Mexico in the 1990s: The main cross-sectional facts." *Review of Economic Dynamics* 13 (1): 238–264.

Brooks, Leah, and Byron Lutz. 2019. "Vestiges of transit: Urban persistence at a microscale." *Review of Economics and Statistics* 101 (3): 385–399.

Bryan, Gharad, Edward Glaeser, and Nick Tsivanidis. 2020. "Cities in the developing world." *Annual Review of Economics* 12:273–297.

Butts, Kyle. 2023. "JUE Insight: Difference-in-differences with geocoded microdata." *Journal of Urban Economics* 133:103493.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro HC Sant'Anna. 2024. *Difference-in-differences with a continuous treatment.* Technical report. National Bureau of Economic Research.

Chen, Yihsu, and Alexander Whalley. 2012. "Green infrastructure: The effects of urban rail transit on air quality." *American Economic Journal: Economic Policy* 4 (1): 58–97.

Conway, Matthew Wigginton, Andrew Byrd, and Michael Van Eggermond. 2018. "Accounting for uncertainty and variation in accessibility metrics for public transport sketch planning." *Journal of Transport and Land Use* 11 (1): 541–558.

Conwell, Lucas. 2023. "Subways or minibuses? Privatized provision of public transit."

Conwell, Lucas, Fabian Eckert, and Ahmed Mushfiq Mobarak. 2023. "Transportation Infrastructure and City-Center Accessibility in the US and Europe."

Davis, Lucas. 2008. "The effect of driving restrictions on air quality in Mexico City." *Journal of Political Economy* 116 (1): 38–81.

Davis, Lucas W. 2021. "Estimating the price elasticity of demand for subways: Evidence from Mexico." *Regional Science and Urban Economics* 87:103651.

Davis, Lucas W, and Paul J Gertler. 2015. "Contribution of air conditioning adoption to future energy use under global warming." *Proceedings of the National Academy of Sciences* 112 (19): 5962–5967.

De Chaisemartin, Clément, and Xavier d'Haultfœuille. 2023. "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey." *The Econometrics Journal* 26 (3): C1–C30.

Deacon, Robert T, and Jon Sonstelie. 1985. "Rationing by waiting and the value of time: Results from a natural experiment." *Journal of Political Economy* 93 (4): 627–647.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review* 106 (3): 479–524.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, geography, and trade." *Econometrica* 70 (5): 1741–1779.

ECLAC. 2015. "Social discount rate (SDR) of Mexico (Spanish) | Regional Observatory on Planning for Development." https://observatorioplanificacion.cepal.org/en/regulatory-frameworks/social-discount-rate-sdr-mexico-spanish.

Fajgelbaum, Pablo D, and Cecile Gaubert. 2020. "Optimal spatial policies, geography, and sorting." *The Quarterly Journal of Economics* 135 (2): 959–1036.

Firth, David. 1993. "Bias reduction of maximum likelihood estimates." *Biometrika* 80 (1): 27–38.

Gaduh, Arya, Tadeja Gračner, and Alexander D Rothenberg. 2022. "Life in the slow lane: Unintended consequences of public transit in Jakarta." *Journal of Urban Economics* 128:103411.

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

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.

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.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics* 225 (2): 254–277.

Guerra, Erick. 2015. "The geography of car ownership in Mexico City: A joint model of households' residential location and car ownership decisions." *Journal of Transport Geography* 43:171–180.

Gupta, Arpit, Stijn Van Nieuwerburgh, and Constantine Kontokosta. 2022. "Take the Q train: Value capture of public infrastructure projects." *Journal of Urban Economics* 129:103422.

Heilmann, Kilian. 2018. "Transit access and neighborhood segregation: Evidence from the Dallas light rail system." *Regional Science and Urban Economics* 73:237–250.

Hernández-Cortés, Danae, Paulina Oliva, and Christopher Severen. 2024. "Comparing Travel Responses to Subway and Bus Rapid Transit: Evidence from Mexico City."

Iniguez-Montiel, Alberto Javier. 2014. "Growth with equity for the development of Mexico: Poverty, inequality, and economic growth (1992–2008)." *World Development* 59:313–326.

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.

Kreindler, Gabriel, Arya Gaduh, Tilman Graff, Rema Hanna, and Benjamin A Olken. 2023. *Optimal Public Transportation Networks: Evidence from the World's Largest Bus Rapid Transit System in Jakarta.* Technical report. National Bureau of Economic Research.

Lee, Kwok Hao, and Brandon Joel Tan. 2024. "Urban transit infrastructure and inequality." *Review of Economics and Statistics*, 1–46.

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.

LeRoy, Stephen F, and Jon Sonstelie. 1983. "Paradise lost and regained: Transportation innovation, income, and residential location." *Journal of Urban Economics* 13 (1): 67–89.

Lind, Nelson, and Natalia Ramondo. 2023. "Trade with correlation." *American Economic Review* 113 (2): 317–353.

McFadden, Daniel. 1974. "The measurement of urban travel demand." *Journal of Public Economics* 3 (4): 303–328.

McMillen, Daniel P, and John McDonald. 2004. "Reaction of house prices to a new rapid transit line: Chicago's midway line, 1983–1999." *Real Estate Economics* 32 (3): 463–486.

Miyauchi, Yuhei, Kentaro Nakajima, and Stephen J Redding. 2021. *The economics of spatial mobility: Theory and evidence using smartphone data.* Technical report. National Bureau of Economic Research.

Monkkonen, Paavo, M. Paloma Giottonini, and Andre Comandon. 2021. " Socioeconomic Segregation in Mexico City: Scale, Social Classes, and the Primate City." In *Urban Socio-Economic Segregation and Income Inequality, A Global Perspective,* edited by Maarten van Ham, Tiit Tammaru, Rūta Ubarevičienė, and Heleen Janssen. The Urban Book Series. Springer.

Oliva, Paulina. 2015. "Environmental regulations and corruption: Automobile emissions in Mexico City." *Journal of Political Economy* 123 (3): 686–724.

Oliva, Paulina, B Kelsey Jack, Samuel Bell, Elizabeth Mettetal, and Christopher Severen. 2020. "Technology adoption under uncertainty: Take-up and subsequent investment in Zambia." *Review of Economics and Statistics* 102 (3): 617–632.

Papke, Leslie E, and Jeffrey M Wooldridge. 2008. "Panel data methods for fractional response variables with an application to test pass rates." *Journal of Econometrics* 145 (1-2): 121–133.

Pollmann, Michael. 2023. "Causal inference for spatial treatments." *arXiv preprint arXiv:2011.00373.*

Redding, Stephen J. 2023. "Quantitative urban models: from theory to data." *Journal of Economic Perspectives* 37 (2): 75–98.

Redding, Stephen J, and Esteban Rossi-Hansberg. 2017. "Quantitative spatial economics." *Annual Review of Economics* 9 (1): 21–58.

Rothenberg, A., and J. Du. 2023. "The Benefits and Costs of Motorization: Evidence from Jakarta."

Samaniego de la Parra, Brenda, and León Fernández Bujanda. 2024. "Increasing the cost of informal employment: Evidence from Mexico." *American Economic Journal: Applied Economics* 16 (1): 377–411.

Severen, Christopher. 2023. "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification." *Review of Economics and Statistics* 105 (5): 1073–1091.

Severen, Christopher, and Arthur A Van Benthem. 2022. "Formative experiences and the price of gasoline." *American Economic Journal: Applied Economics* 14 (2): 256–284.

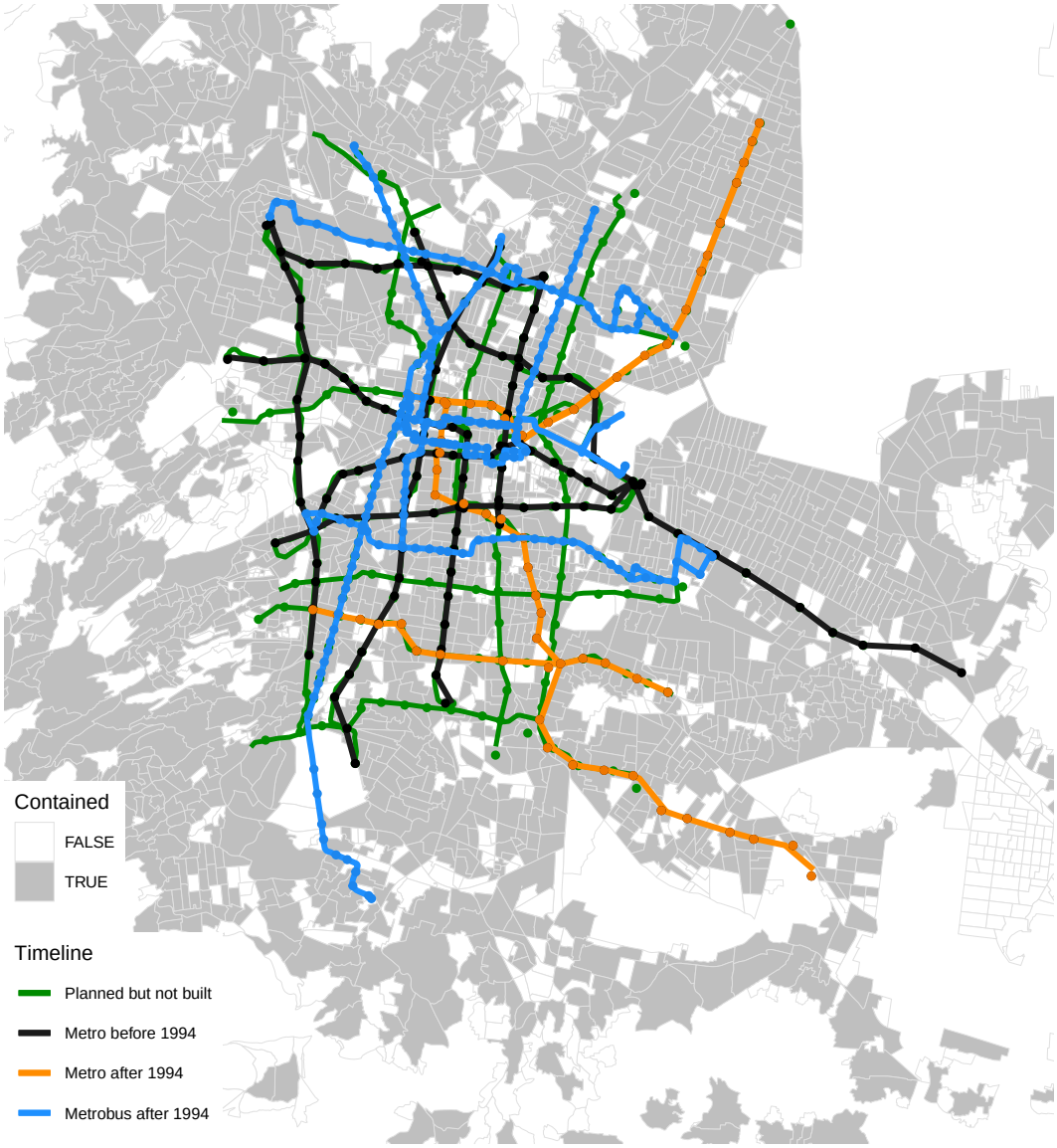
Train, Kenneth E. 2009. *Discrete choice methods with simulation.* Cambridge University Press.

Tsivanidis, Nick. 2023. "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio."

Zárate, David R. 2023. *Spatial Misallocation, Informality and Transit Improvements: Evidence from Mexico City.* Technical report.

Zhou, Yiyi, and Shanjun Li. 2018. "Dynamics of Technology Adoption and Critical Mass in the Market for Electric Vehicles." *Journal of Industrial Economics* 2:423–480.

Figure 1: Built and planned rapid transit lines in Greater Mexico City by 2017.



This map shows the central portion of Mexico City and planned an built rapid transit lines. Subway (Metro) lines built before 1994 are in black and those built after 1994 are in orange. BRT (Metrobús) lines built between 1994 and 2017 are in blue. Planned lines present in the original transit expansion plan, *Plan Maestro 1985*, but that were not built are shown in green. We use these planned-but-not-built lines to construct a control group in some specifications (see Section 3.1.2). Outlined areas are 1994-basis census tracts, and are filled with gray if they are observed in all three periods.

Figure 2: Mode Choice Neseted Logit Structure

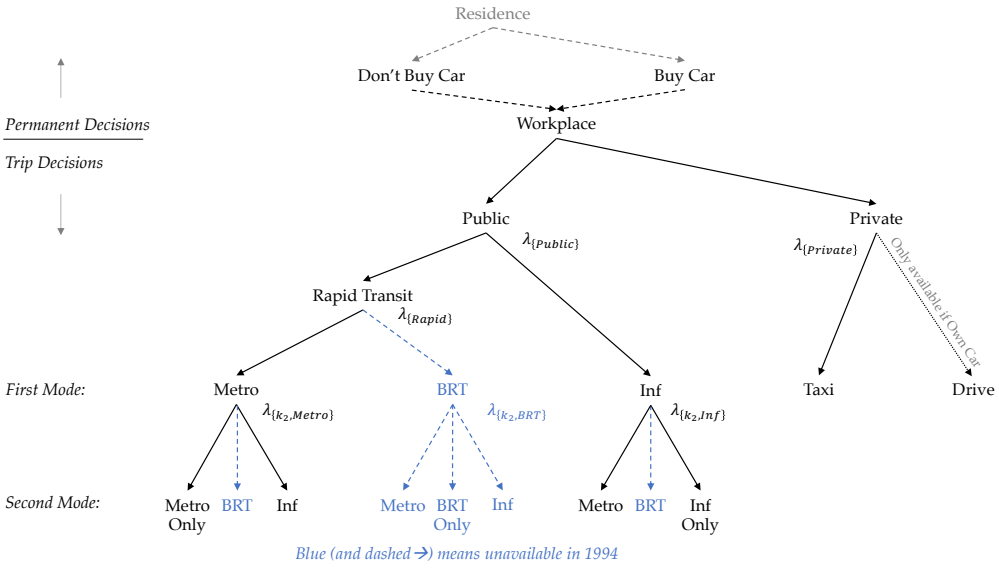


Figure 3: Market Access and Vehicle Choice (Semi-Elasticity)

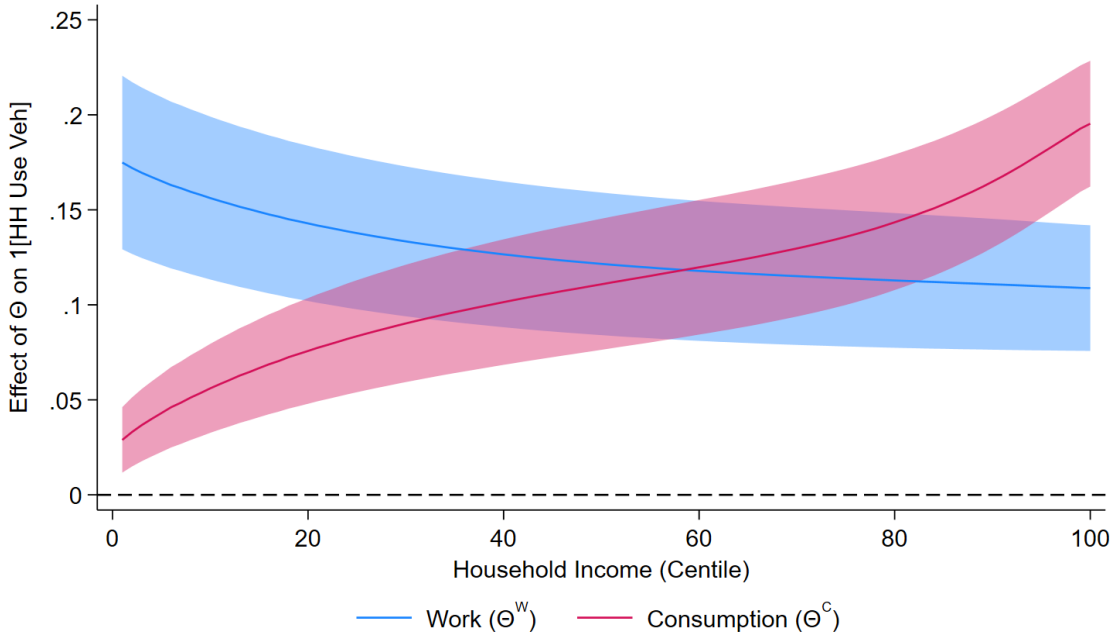
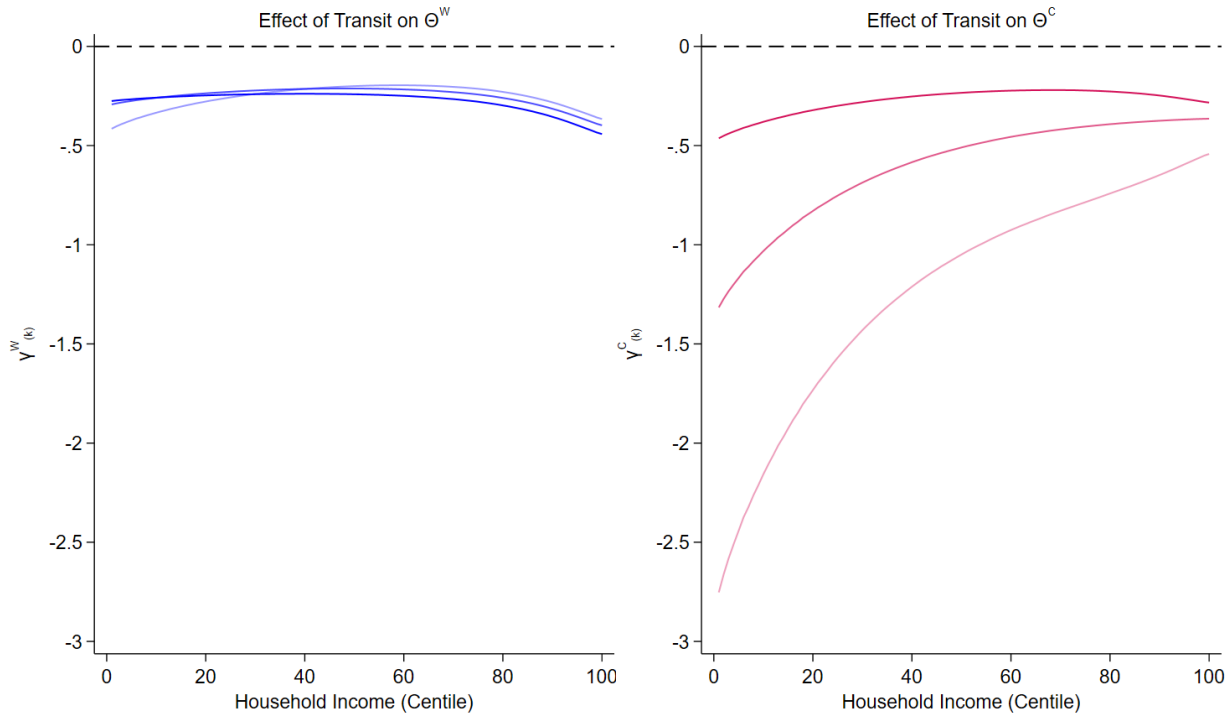


Figure 4: Effect of Transit on Relative Market Access by Income



asdf

Figure 5: Imputed Effect of Transit on Car Ownership via Relative Market Access

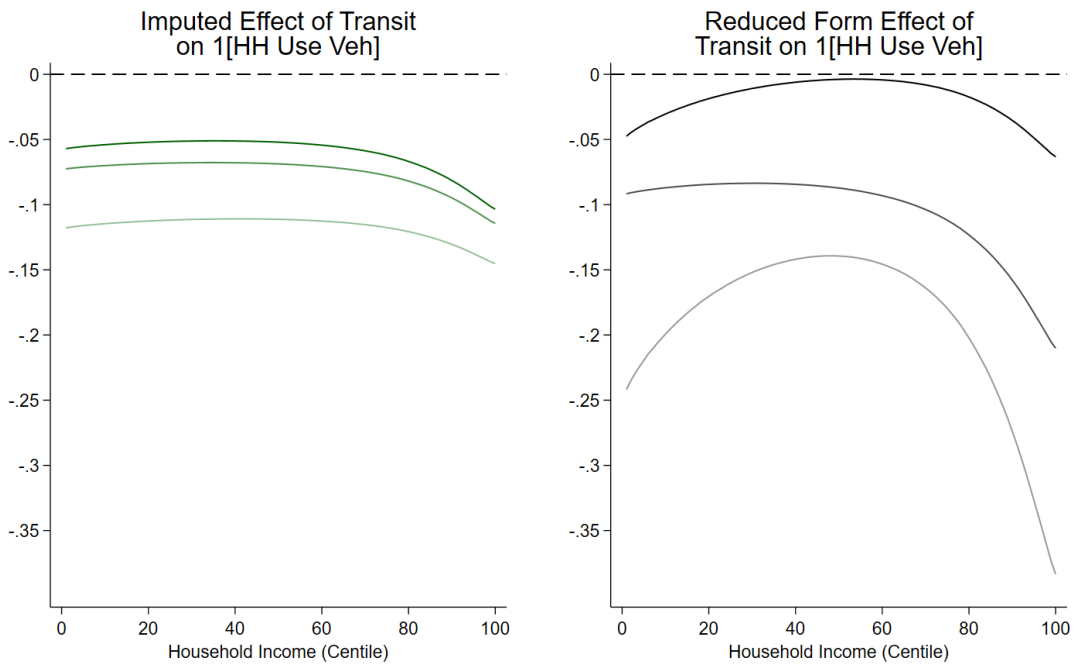


Table 1: Summary Statistics, EOD Data

	1994			2007			2017		
	Trip Share	Ave. Dist. (km)	Ave. Time	Trip Share	Ave. Dist. (km)	Ave. Time	Trip Share	Ave. Dist. (km)	Ave. Time
Panel A. Mode Use Summary and Travel Time									
All		8.5	46.9		8.7	51.6		8.6	54.5
Transit	0.748	8.9	51.2	0.656	9.4	58.9	0.635	9.4	63.3
Metro (subway)	0.215	13.9	70.6	0.216	14.5	79.4	0.225	13.9	81.6
Metrobús (BRT)	–	–	–	0.011	12.6	73.3	0.055	11.1	71.3
Bus/Colectivo	0.723	9.0	51.7	0.617	9.5	59.6	0.550	9.5	64.3
Private vehicle	0.225	7.9	35.5	0.285	7.9	40.9	0.308	7.9	42.5
Taxi	0.029	4.7	26.0	0.067	5.2	31.3	0.067	4.9	30.9
No. of Trips		48397			73328			63481	
Panel B. Multi-modal Trips									
Trips with 1 mode	0.808			0.811			0.806		
Trips with 2 modes	0.191			0.181			0.184		
Trips with 3+ modes	0.001			0.008			0.010		
Panel C. Household Vehicles									
	1994			2007			2017		
	Share Using Veh.	Share Owning Veh.	Ave. No. of Veh.	Share Using Veh.	Share Owning Veh.	Ave. No. of Veh.	Share Using Veh.	Share Owning Veh.	Ave. No. of Veh.
By Household	0.318	–	–	0.364	0.453	0.61	0.387	0.440	0.61
No. of HHs	21733			33546	39234		35977	42602	

Sample includes households residing in census tracts observed in all three survey years and their trips that originate from home. Each trip may consist of multiple modes, so shares do not add to one. Walking-only and biking-only trips are excluded, and walking and biking do not count toward additional modes. Travel times for individual trips are top-coded to 240 minutes. Data on vehicles owned by households are not available in 1994. Sample sizes for share using vehicles are conditioned on observing any trips in the household. Observations weighted by sample weights.

Table 2: Scale Effects in Household Vehicle Use (Baseline and Muni Designs)

	1[Any HH Vehicle Use]			
	(1)	(2)	(3)	(4)
1(+) Line(s)	-0.038**	-0.030*	-0.025	-0.022
	(0.014)	(0.014)	(0.016)	(0.016)
2 Lines		-0.108***		-0.052*
		(0.022)		(0.024)
3+ Lines		-0.114***		-0.036
		(0.027)		(0.033)
Cumulative Eff.		-0.253***		-0.110**
		(0.032)		(0.042)
Pr(CRS of 2nd Line)		[0.007]		[0.344]
Pr(CRS of 3rd+ Line)		[0.160]		[0.968]
Design	Baseline	Baseline	Muni	Muni
Tract FEs	Y	Y	Y	Y
Year FEs	Y	-	Y	-
Muni×Year FEs	-	Y	-	Y

91256 observations. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table 3: Scale Effects in Household Vehicle Use (PM85 and Scaffolded Designs)

	1[Any HH Veh. Use]		
	(1)	(2)	(3)
1st (1+) Line | 0 Lines Present	-0.014	-0.048**	-0.029+
	(0.017)	(0.015)	(0.017)
2nd Line | 1 Line Present		-0.080**	-0.068+
		(0.027)	(0.035)
3rd Line | 2 Lines Present		-0.118*	-0.088
		(0.048)	(0.057)
Cumulative Eff.		-0.246***	-0.186**
		(0.057)	(0.068)
Pr(CRS of 2nd Line)		[0.302]	[0.325]
Pr(CRS of 3rd+ Line)		[0.289]	[0.509]
Design	PM85	Scaffold	Scaffold
Tract FEs	Y	-	-
Tract×Scaffold FEs	-	Y	Y
Year FEs	Y	-	-
Year×Scaffold FEs	-	Y	-
Muni×Year×Scaffold FEs	-	-	Y
N	39114	90499	90499

UPDATE

Table 4: Scale Effects in Household Vehicle Use (PM85 Design)

	1[Any HH Veh. Use]			
	(1)	(2)	(3)	(4)
1st Line	0.015	-0.033	-0.025	-0.031
	(0.031)	(0.021)	(0.016)	(0.021)
2nd Line	-0.094***	-0.050	-0.066*	-0.046
	(0.025)	(0.034)	(0.026)	(0.035)
3rd+ Line	-0.074*	-0.054	-0.090*	-0.064
	(0.030)	(0.045)	(0.037)	(0.047)
Cum. Effect (2 Lines)	-0.079*	-0.083*	-0.091**	-0.077+
	(0.036)	(0.039)	(0.028)	(0.041)
Pr(CRS of 2nd Line)	[0.011]	[0.683]	[0.204]	[0.710]
Cum. Effect (3+ Lines)	-0.153***	-0.137*	-0.181***	-0.141*
	(0.045)	(0.058)	(0.047)	(0.063)
Pr(CRS of 3rd+ Line)	[0.337]	[0.797]	[0.266]	[0.615]
Design	PM85	PM85	PM85	PM85
Tract FEs	Y	-	Y	Y
Tract×Stack FEs	-	Y	-	-
Year FEs	Y	-	Y	-
Year×Stack FEs	-	Y	-	Y
Only PM85 Sample	Y	Y	-	-
Only PM85 Variation	-	Y	Y	Y
Stacked Estimation	-	Y	-	-
N	43245	45694	91256	91256

UPDATE

Table 5: Model Estimates – Elasticities

	Single Class		Latent Class		Latent Class	
	Work $\pi^W_{j\|iat}$	Cons. $\pi^C_{j\|iat}$	Work $\pi^W_{j\|iatg}$	Cons. $\pi^C_{j\|iatg}$	Work $\pi^W_{j\|iatg}$	Cons. $\pi^C_{j\|iatg}$
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln d^c_{iajt}$ $(= \epsilon^c)$	3.602***	4.991***	2.712***	4.506***		
	(0.021)	(0.039)	(0.017)	(0.037)		
$\ln d^c_{iajt,g=\text{lo}}$ $(= \epsilon^c_{g=\text{lo}})$					3.040***	8.553***
					(0.021)	(0.064)
$\ln d^c_{iajt,g=\text{mid}}$ $(= \epsilon^c_{g=\text{mid}})$					3.584***	5.146***
					(0.022)	(0.044)
$\ln d^c_{iajt,g=\text{hi}}$ $(= \epsilon^c_{g=\text{hi}})$					2.068***	3.050***
					(0.013)	(0.024)
N	2521092	2345527	7563276	7036581	7563276	7036581

Estimated by PPMLHDFE. Standard errors clustered by residential tract, + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 6: Model Estimates and Evaluation, Single Class Model

	Θ^W_{ij}	Θ^C_i	1[Any HH Veh. Use]		
				Imp.	Est.
	(1)	(2)	(3)	(4)	(5)
Θ^W_{ij} $(= \mathcal{B}^W)$			0.201***		
			(0.028)		
Θ^C_i $(= \mathcal{B}^C)$			0.076***		
			(0.017)		
1 Line $(= \gamma^c_{(1)})$	-0.281***	-0.244***		-0.075	-0.018
	(0.012)	(0.011)			(0.015)
2 Lines $(= \gamma^c_{(2)})$	-0.258***	-0.471***		-0.088	-0.130***
	(0.014)	(0.019)			(0.025)
3+ Lines $(= \gamma^c_{(3)})$	-0.261***	-0.946***		-0.125	-0.245***
	(0.016)	(0.029)			(0.033)
Tract FEs	Y	-	-	-	-
Tract×1[Work] FEs	-	Y	Y	-	Y
Year FEs	Y	Y	Y	-	Y
N	61164	85668	85668		85668

Standard errors clustered by residential tract, + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Appendices and Supplemental Results

to accompany

TITLE

by

AUTHORS

November 2023

Appendix

A Data Description, Construction, and Details **A-2**
A.1 Data Sources . . . A-2
A.1.1 Origin-Destination Surveys (EODs) . . . A-2
A.1.2 Census Data . . . A-2
A.1.3 Infrastructure Data . . . A-3
A.2 Data Standardization and Normalization . . . A-3
A.2.1 Geonormalization . . . A-3
A.3 *Plan Maestro 1985* . . . A-4

B Travel Time Simulation and Historically Reasonable Travel Times **A-5**
B.1 Approximating Historical Infrastructure . . . A-6
B.2 Routing Engine and Time Simulation . . . A-7

C Wage Prediction and Normalization **A-9**
C.1 Wage Prediction . . . A-9
C.2 Comparison of EOD Wages with ENIGH Wages . . . A-10

D Model Derivations and Extensions **A-11**
D.1 Derivations . . . A-11
D.1.1 Equation (8) . . . A-11
D.1.2 Equation (10) . . . A-12
D.1.3 $\mathcal{B}$ recovers β . . . A-13
D.2 Endogenizing Residential Choice . . . A-13
D.2.1 With Observable Rents . . . A-14
D.2.2 Without Observable Rents . . . A-14

E Additional Effects of Transit on Automobile Adoption and Travel Behavior **A-15**
E.1 Distance Bins and $\bar{d}$ . . . A-15
E.2 Returns to Scale for Subway versus BRT . . . A-15
E.3 Up to 5 Lines of Treatment and Comparisons with Discrete Treatment . . . A-16

A Data Description, Construction, and Details

A.1 Data Sources

A.1.1 Origin-Destination Surveys (EODs)

Data come from several sources. Our primary sources for household and and travel data are three travel surveys from Greater Mexico City. These surveys cover households in the 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 briefly detail the sourcing of each wave:

EOD 1994: We obtained a copy of EOD 1994 from the Secretaía de Movilidad de la CDMX (SEMOVI).[53] The data consist of three tables, one each for households, people, and trips. Packaged with the data are a short codebook as well as a scan of the original survey instrument. From these materials, we deduced that several variables from the survey form were not input into the data (such are car ownership). We have been unable to determine if the raw (i.e., paper) data or less post-processed versions of the data exist.

Unfortunately, we have also been unable to determine the precise geography used to for EOD 1994, in large part because geographies used in EOD 1994 do not precisely INEGI's typical scheme for numbering small geographic units. While we can concord many geographies (see below), there are some geographies which we cannot link.

EOD 2007: These data were obtained in 2009 via a freedom of information (Transperencia) request from the precursor of SEMOVI (which was called Oficina de Información Pública de la Secretaria de Transportes Y Vialidad del Distrito Federal).

EOD 2017: These publicly available data were obtained from INEGI. We augment this data with a confidential merge on household location and survey day, conducted by INEGI at our request.

A.1.2 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.

53. When we initially contacted SEMOVI, they were unaware that they possessed this file. We obtained a subset of this data from Erick Guerra (e.g., Guerra 2015).

A.1.3 Infrastructure Data

Our baseline data for subway comes from a conversion of Mexico City's KML files by transit writer Jose Gonzalez.[54] At the time we began the project, this was the only shapefile offering full coverage of Mexico City's subway system. We validated this data by comparing it to Google Maps and diagrammatic maps of the subway system. We directly converted data on Metrobús stations from official KML files to shapefiles.[55] We supplement this data with information on station openings that we primarily retrieve from the Spanish- and English-language Wikipedia articles about the system, and verify with additional search.

A.2 Data Standardization and Normalization

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 similar to census tracts 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 three caveats. First, EOD 1994 uses a modified AGEB layer in which very small AGEBs are unioned together instead of standard AGEBs. We received shapefiles for these modified AGEBs from Erick Guerra, as neither we nor INEGI can find the originals. Second, metadata containing information about the geographic projection used for the EOD 1994 and EOD 2007 maps was not available with the data or was incorrect. We therefore rubbersheeted these projections to the 2017 standard (for which we have projection metadata) by hand using ArcGIS. Finally, districts are not a standard INEGI geography; we form them from AGEBs when possible. In EOD 2017, we do not know origin and destination AGEB (we only know origin and destination districts, in addition to residential AGEB), so we use the shapefile of 2017 districts included with the data.

A.2.1 Geonormalization

We standardize residential data to reflect 1990 EOD AGEBs and travel (origin-destination) data to reflect 2017 travel districts. When an observation cannot be uniquely assigned to a single normalized AGEB or travel district, multiple observations are created with areal weights that sum to one. An observation-specific identification code is retained across these splits to permit adjusting standard errors to reflect induced cross-location error correlation.

Overall, we can consistently match (2000–2500) 1990 EOD AGEBs across all 1994, 2007, and 2017. We can also match around 140 travel districts over time (generating 19,600 OD pairs). For some travel data (trips beginning or ending at home), we know the AGEB of residence and combine that with the origin or destination district.

54. See this link for a description; the files can be downloaded at https://github.com/josecarlosgonz/mexicoCityMetro.

55. Modern data can be found at https://datos.cdmx.gob.mx/dataset/geolocalizacion-metrobus.

Data and Sample Harmonization

The below details the challenges in combining EOD data across waves and combining the 2017 EOD with 2015 Conteo data for income prediction.

EOD Data Harmonization

	1994 EOD	2007 EOD	2015 Conteo	2017 EOD	94-07-17 EOD harmonized
Sample					
People:	If trips≥1	All	All	All	If trips≥1
Trips:	Non-walking, Weekday	Non-walking, Weekday	- -	All Weekday/end	Non-walking Weekday
Geography					
Residences:	1990 AGEB	2007 AGEB	2015 AGEB	2015 AGEB	1990 AGEB
Trips (O-D):	1990 AGEB	2007 AGEB	-	2017 District	2017 District
Automobile					
Car in household	-[56]	Yes	Yes	Yes	-
Car used for trip	Yes	Yes	-	Yes	Yes
Demographics					
Income	Yes	Yes	Yes	-	Predicted
Education	-	Yes	Yes	Yes	-
Occupation/Industry	Yes (5/0)	Yes (8/8)	Text	-	-
Age/Sex/HH Status	Yes	Yes	Yes	Yes	Yes

A.3 *Plan Maestro 1985*

We construct a counterfactual subway network according to the *Programa Maestro del Metro, Segunda Revision, Version 1985*, which we refer to more colloquially as the *Plan Maestro 1985* (PM85). The PM85 calls for a fifteen line buildout of the Metro that looks relatively similar to the state of the combined Metro and Metrobús system as of 2017. There is some evidence of substitution from subway to BRT along some corridors, so it seems reasonable to use the PM85 as a counterfactual. Also see Zárate (2023) for a discussion of the PM85.

The starting point is the map on plat page V-65, which contains both built and planned lines and stations, along with some annotations. Figure A1 shows this plat; the thicker lines are those built before 1985, the thinner lines are those proposed to be built afterwards. Of these, Line 9 and Line A were completed before our study period, but were built as shown in this map. We rubber-sheet this map based on already constructed lines to our data on Mexico City subway stations. We focus on station locations, and where station locations are unclear, we use detail plats in provided in the rest of the PM85 to infer the most likely location.

56. The 1994 EOD includes a question about car availability in the household, but it does not appear to have been digitized.

B Travel Time Simulation and Historically Reasonable Travel Times

Our mode choice model requires the full distribution of travel times that agents face. We do not observe the full distribution of travel times across all possible origin-destination pairs and all modes in all three of our survey years. The default way of obtaining this data (via queries to Google Maps or OSM) does not allow for historical sampling. While one option would be to use modern data and assume its validity in the past, such an approach misses several key features. First, road infrastructure may have changed, shifting the utility of automobile transport. Second, fixed transit (such as subways and bus rapid transit) typically replaces other transit modes, such as buses. Third, travel times conditional on mode may have shifted, both as a result of changes in service or infrastructure, or as a result of changes in congestion.

We seek to approximate historically reasonable travel times by reconstructing approximations of the infrastructure and transit services environments to the survey years under study. Specifically, this consists of:

1. Approximating year-specific road and transit networks.
2. Equipping a routing engine with the networks in step 1) to simulate uncongested travel times for each mode for a broad sample of potential origins and destinations for each survey year.
3. Allowing the model to fit variable congestion across years.

We detail each of these tasks in following subsections.

To fit our model, we require travel times between all locations on all modes and all combinations of modes in each of years. This is a large problem, complicated by the fact that we do not observe historical transit service and non-rapid transit is only partially observed. Our solutions involve:

- To accommodate multi-modal trips, we first use the implementation the Rapid Realistic Routing on Real-world and Reimagined (R5) networks implemented by Conveyal. This accommodates some multi-modal travel and accounts for expected waits due to line or mode transfers.
- However, because we need to simulate travel times between every location using *every* mode combination, sometimes we constrain the model and combine modes to form routes ex post.
- We backcast the subway and BRT service starting with a GTFS feed of travel in 2017 and removing stations or lines as appropriate for 1994 and 2007.
- We backcast roadways by removing highway designations to upgraded highways.
- To estimate travel times for buses and informal transit, we combine data from three sources. First, we use modern Mexico City GTFS feeds to provide a set of bus service data. We also use archived data from Mapatón and ViaDF, two services which asked riders to map transit routes between 2010 and 2015.

More details are provided in the Appendix.

B.1 Approximating Historical Infrastructure

We lack precise data on historical road and transit networks, so we make a number of assumptions to develop reasonable approximations to historical networks that, when aggregated, provide reasonable historical travel times. Our primary method is to backcast modern data on transit service to historical periods.

Approximate Historical Transit Service

The modern standard for communicating transit system information, the General Transit Feed Specification (GTFS), is available but incomplete in Mexico City for 2017 and unavailable for 1994 and 2007. Feeds from 2020 and earlier omit many keep bus routes, which Mexico City's transit agency finally began including all coverage of in 2021.[57]

Rapid transit in Mexico City earlier GTFS feeds appears to be fully covered. So, for rapid transit, we begin with the official March 21, 2017 transit feed, as published by the Mexico City government and archived online by OpenMobilityData [58]. We backcast this data to 2007 by removing all Métrobus service excepting the first 37 stations of Line 1, as well as Metro Line 12. For Métrobus Line 1, we clip service to the first 37 stations, assuming routes run with the same frequency as they did in 2017. We further backcast to 1994 by fully removing Métrobus Line 1 and Metro Line 8 and Metro Line B.

For non-rapid modes, recreating historical service networks is more challenging due to a lack of official data. Our approach consists of two primary components. We backcast recent data on bus service to historical periods. We then combine this data with crowd-sourced approximations of historical formal bus and informal transit networks.

For the first of these components (backcasting bus service to historical periods), we begin with the official October 27, 2021 transit feed as our baseline datasource. We first verify that for rapid modes, service frequency was similar to available 2017 feeds.[59] This supports using data from 2021 to approximate service in 2017. This GTFS feed includes data on trolleybus, light rail, Mexicable, and a new commuter rail service. We exclude Mexicable and the commuter rail service, as they were not yet available in 2017. We then assume that this was the service level available by formal non-rapid modes between 1994 and 2017.[60]

To accommodate informal transit, we first turn to two informal data sources: Mapatón and ViaDF. Mapatón was a hackathon-style event conducted in 2016 intended to capture the extent of informal transit in Mexico City.[61] A team of developers created an mobile-based app that volunteers could use to track and geocode informal transit service. To incentive coverage in places where fewer people had mobile phones, the developers created a prize scheme that over-rewarded

57. Historical transit feeds are found at https://transitfeeds.com/p/mexico-city-federal-district-government/70/20170321 the most recent feed is available at https://datos.cdmx.gob.mx/dataset/gtfs.

58. Source: https://transitfeeds.com/p/mexico-city-federal-district-government/70/20170321)

59. COVID-19 resulted in reduced transit service in many cities, but for Mexico City rapid transit modes, service appeared approximately equivalent.

60. This is unlikely to be exactly correct, but hopefully gives a reasonable approximation of non-rapid transit service, which tends to be more uniform in its coverage.

61. See, e.g., https://nextcity.org/urbanist-news/big-data-mexico-city-mapathon-gamifies-crowdsourcing for description. The official Mapatón website is no longer available, but can be found here: https://web.archive.org/web/20180329050356/http://www.mapatoncd.mx/.

rare routes. We use the shapefile-based version of the data.[62]

ViaDF was also a crowdsourced project intended to provide mapping of informal transit in Mexico City that was primarily active from 2010-2015.[63] This website provided a way for users to log their routes for others to use.

We hired a programmer with GTFS experience to convert the Mapatón and ViaDF data into a GTFS feed compatible with the existing official GTFS feeds. Because service frequency and speed is only available for a subset of these data, the simualted GTFS feed was create assuming a stop density of 330m and travel speed of 14kph.

Approximate Historical Road Networks

To the best of our knowledge, there are no digitally formatted historical road networks available for Mexico City in 1994 or 2007, while there have not been substantial changes between 2017 and today. We are primarily concerned with major shifts in the road infrastructure that would have greatly increased the potential speed an automobile can take. In a word, highways. Between 1994 and 2017, Mexico City and the State of Mexico greatly expanded the outer beltway that circles Mexico City, referred to as the Anillo Periférico, or just Periférico. This was the primary substantial change in automoblity during this period.

We approximate the historical network using parsimonious adjustments to the modern network. We begin with the Open Street Map (OSM) data of modern road network of Mexico City and the surrounding region as depicted in Open Street Maps. We then downgrade the portions of the Periférico that were constructed after 1994 and after 2007 to function as surface streets. This reduces the speed limit that the routing engine assigns to these streets. Specifically, we downgrade from OSM highway class 'motorway' to OSM highway class 'secondary'. The routing engine we use sets defaults speeds for these road classes to 65mph and 35mph, respectively. All told, this amounted to 144 changes for 2007 and 407 changes for 1994.[64]

Specifically, the workflow consists of:

1. Obtain and download OSM data from Open Street Map.
2. Edit the highway class data in QGIS. To determine which road segments should be downgraded, we overly data on road constructions dates from Mexico City (DATA SOURCE) and downgrade those that were constructed each of our reference years.
3. Export the edit file as a .geojson files.
4. Import the .geojson files to JOSM and reformat as .osm.pbf files (the required format for the routing engine).

B.2 Routing Engine and Time Simulation

We combine the street network data and GTFS feeds into three vintages, and then upload theserouting servers hosted by Conveyal to run the R5 routing algorithm (Conway, Byrd, and Van

62. Formerly available here: https://web.archive.org/web/20161219021755/http://blog.mapatoncd.mx/basededatos/

63. Website and data are here: https://viadf.mx/.

64. We have available a list of all the individual changes that were made. An alternative would have been to simply remove the appropriate sections of the Periférico. However, typically highways, particularly urban highways, replace existing roads, so this approach seems reasonable.

Eggermond 2018).[65] Conveyal's engine uses the R5 routing algorithm (R5 is "Rapid Realistic Routing on Real-world and Reimagined networks"), and is available as an open source application.[66] We use the server application because of the large scale of the simulations we run.

We create a sample of origins and destinations in order to approximate travel times. For tract-level locations in Mexico City, we use block-level data on built area to create tract-level weighted centroids as origin points. Outside of Mexico City, we use simple tract centroids as origin points. For district-level locations (which are aggregated from tracts), we sample 18 points from each districts according to built-area weights within Mexico city, and by drawing from areas covered by valid tracts outside of Mexico City. This gives us a frame of 4,139 origin points and 3,492 destination points. Denote these as $\mathcal{O}$ and $\mathcal{D}$, respectively.

For walking, driving, and informal modes, we using the routing engine to calculates times for each pair in $\mathcal{O} \times \mathcal{D}$. To force the routing engine to include formal modes even when it was not optimal, we create a layer of enter and exit points to the formal (rapid) transit system, denoted $\mathcal{F}$. We then calculated the times for each of $\mathcal{R}_1 = \mathcal{O} \times \mathcal{F}$, $\mathcal{R}_2 = \mathcal{F} \times \mathcal{F}$, and $\mathcal{R}_3 = \mathcal{F} \times \mathcal{D}$. We find the values that minimize travel time between each pair $\mathcal{O} \times \mathcal{D}$ as mediate by travel through $\mathcal{F} \times \mathcal{F}$. We then take the average of all $\mathcal{O} \times \mathcal{D}$ across sampling points with each district.

65. See https://conveyal.com/.
66. See https://github.com/conveyal/r5.

C Wage Prediction and Normalization

C.1 Wage Prediction

We create a measure of predicted wage conditional on residential location to examine heterogeneous responses to treatment by income and to inform latent class selection in the structural model. The 1994 and 2007 EODs contain reported wage and wage frequency. However, the 2017 EOD does not include wages or income. To overcome the lack of wage data, we predict wages using the 2015 conteo and worker characteristics observed in both the 2017 EOD and 2015 conteo.[67]

Neither of the public versions of the 2017 EOD nor the 2015 conteo contain fine- (AGEB-)level residential location, and so linking geographic locations in the two data sources is challenging. Although Mexico City is not as segregated as, for example, many US cities, residential location can reflect useful information about wage potential (Monkkonen, Giottonini, and Comandon 2021). We overcome this by confidential versions of both datasets that include residential tract (AGEB). We then use residential tract as one of the predictors of wage.

Our first step is harmonize wages across time and reporting frequency. We adjust nominal income to 2015 prices using the "Consumer Price Index: All Items: Total for Mexico" series and adjust 1994 income by pay period to monthly income.[68] We implement a common top-code wage of 200,000 pesos/month.

Let w_{ic} be the real monthly income for worker i residing in tract c. We estimate a Mincer-style wage model that includes residential location for each year. That is, we estimate:

$$\ln(w_{ict}) = \alpha_{ct} + X_{it}\lambda_t + e_{it}$$

for each $t = 1994, 2007, 2015$. We then predict $\omega_{it|c} \equiv \widehat{\ln(w_{ict})} = \hat{\alpha}_{ct} + X_{it}\hat{\lambda}_t$ for $t = 1994, 2007$. For 2017, we use estimates from 2015 data harmonized to 2017 covariates to predict wage: $\widehat{\ln(w_{ic2017})} = \hat{\alpha}_{c2015} + X_{i2017}\hat{\lambda}_{2015}$.

Because of variation in the survey data across years, the contents of X vary a bit in each survey wave:

- 1994: age, sex, interaction of age and sex, status in household, and interactions of sex and status in household.
- 2007 & 2017: age, sex, interaction of age and sex, education (3 bins), interactions of sex and education, and status in household.

Results

We summarize and evaluate the results of this prediction exercise in several tables and figures. All prices have been converted into 2015 pesos using the OECD's CPI series for Mexico.Table A9 shows that mean observed real personal income was $9,495 MXN in 1994 and $8,009 in 2007 MXN. We found this decline surprising given that real GDP in Mexico was higher in 2007 than 1994. However, Binelli and Attanasio 2010 find that real wages fell substantially following the Peso

67. The conteo is similar the census but somewhat smaller in scope. Censuses are collected every decade on years ending in 0, and conteos are intercensal surveys collected in years ending in 5.

68. The CPI is available at https://fred.stlouisfed.org/series/MEXCPIALLAINMEI; the original source is OECD – Main Economic Indicators.

Crisis in 1995 and barely crept up through at least 2002. Averaging per capita income across all of urban Mexico, Iniguez-Montiel 2014 find some recovery between 2002 and 2008. Because of the ambiguity, we perform additional analysis using ENIGH and report results both rank-normalized and real income.

Predicted income matches observed income on average (this is unsurprising given our prediction methodology). Predicted personal income is $9,667 MXN in 1994, $8,011 MXN in 2007, and $7,925 MXN in 2017. However, there is some compression of variation in the income distribution, as the standard deviation of predicted wages is 69% of the observed standard deviation in 1994 and 54% in 2007.

Figure A6 shows the distribution of monthly personal income for predicted and observed income in each year. It reveals that the modal income is somewhat higher in predicted income than observed income, and that predicted values smooth some lumpiness in observed values. However, on average, Figure A7 shows a relatively tight correlation between the level of observed income and predicted income in each year. These binscatter plots reveal that the predicted values are, on average, close to a linear function of observed income.

Rank Normalization

Because of the difficulties predicting the variation in income and normalizing levels between 1994 and later years, we provide an alternative delineation of income. We assign each predicted income to its income rank, normalized to the nearest centile between 1 and 100. Figure A8 compares average observed log income to income rank for 1994 and 2017, revealing a near monotonic relationship.

C.2 Comparison of EOD Wages with ENIGH Wages

D Model Derivations and Extensions

D.1 Derivations

D.1.1 Equation (8)

Equation (7) defines ω's flow utility of working in j conditional on choices ia and group membership g as

$$\mathcal{V}^W_{iajg,\omega}\left(\omega^W_{jg}, d^W_{iajg}\right) = \frac{1}{\epsilon^W_g}\left(\ln \omega^W_{jg} + \epsilon^W_g \ln d^W_{iajg} + e^W_{ijg,\omega}\right) + r\left(p\mathbb{E}\left[\max_k \mathcal{V}^W_{iakg,\omega}\left(\omega^W_{kg}, d^W_{iakg}\right)\right] + (1-p)\,\mathcal{V}^W_{iajg,\omega}\left(\omega^W_{jg}, d^W_{iajg}\right)\right).$$

To solve, we first note that the rightmost $\mathcal{V}^W_{iajg,\omega}$ term can immediately be combined with the LHS. Doing this, setting $r^* = \frac{1}{1-r(1-p)}$, and rearranging yields:

$$\mathcal{V}^W_{iajg,\omega}\left(\omega^W_{jg}, d^W_{iajg}\right) = \frac{r^*}{\epsilon^W_g}\left(\ln \omega^W_{jg} + \epsilon^W_g \ln d^W_{iajg} + e^W_{ijg,\omega}\right) + r^* rp\mathbb{E}\left[\max_k \mathcal{V}^W_{iakg,\omega}\left(\omega^W_{kg}, d^W_{iakg}\right)\right].$$

We now take the expected maximum of both sides, i.e.,

$$\begin{aligned}\mathbb{E}\left[\max_{k'} \mathcal{V}^W_{iak'g,\omega}\left(\omega^W_{k'g}, d^W_{iak'g}\right)\right] &= \mathbb{E}\left[\max_{k'} \frac{r^*}{\epsilon^W_g}\left(\ln \omega^W_{k'g} + \epsilon^W_g \ln d^W_{iak'g} + e^W_{ik'g,\omega}\right) \right.\\ &\qquad \left. + r^* rp\mathbb{E}\left[\max_k \mathcal{V}^W_{iakg,\omega}\left(\omega^W_{kg}, d^W_{iakg}\right)\right]\right]\\ &= \mathbb{E}\left[\max_{k'} \frac{r^*}{\epsilon^W_g}\left(\ln \omega^W_{k'g} + \epsilon^W_g \ln d^W_{iak'g} + e^W_{ik'g,\omega}\right)\right]\\ &\qquad + r^* rp\mathbb{E}\left[\max_k \mathcal{V}^W_{iakg,\omega}\left(\omega^W_{kg}, d^W_{iakg}\right)\right],\end{aligned}$$

where the last equality holds because $\mathbb{E}\left[\max_k \mathcal{V}^W_{iakg,\omega}\left(\omega^W_{kg}, d^W_{iakg}\right)\right]$ does not depend on k'. But the only difference between $\mathbb{E}\left[\max_k \mathcal{V}^W_{iakg,\omega}\left(\omega^W_{kg}, d^W_{iakg}\right)\right]$ and the LHS is the indexing, so:

$$\mathbb{E}\left[\max_k \mathcal{V}^W_{iakg,\omega}\left(\omega^W_{kg}, d^W_{iakg}\right)\right] = \frac{r^*/\epsilon^W_g}{(1-r^* rp)}\mathbb{E}\left[\max_k \left(\ln \omega^W_{kg} + \epsilon^W_g \ln d^W_{iakg} + e^W_{ikg,\omega}\right)\right].$$

Because of the EV Type I distributional assumption,

$$\mathbb{E}\left[\max_k \left(\ln \omega^W_{kg} + \epsilon^W_g \ln d^W_{iakg} + e^W_{ikg,\omega}\right)\right] = \ln\left(\sum_k \omega^W_{kg}\left[d^W_{iakg}\right]^{\epsilon^W_g}\right) = \ln \Phi^W_{iag},$$

where the last equality is true by definition.

Bringing these derivations together,

$$\mathcal{V}^W_{iajg,\omega}\left(\varpi^W_{jg}, d^W_{iajg}\right) = \frac{r^*}{\epsilon^W_g}\left(\ln \varpi^W_{jg} + \epsilon^W_g \ln d^W_{iajg} + e^W_{ijg,\omega} + \frac{r^* rp}{(1-r^* rp)} \ln \Phi^W_{iag}\right).$$

Substituting in $r^* = \frac{1}{1-r(1-p)}$ and rearranging yields Equation (8).

D.1.2 Equation (10)

For Equation (9) (the sub-utility for a consumption trip), note that we can rewrite the RHS as

$$\frac{1}{\epsilon^C_g} \ln \varpi^C_{jg} + \ln d^C_{iajg} + \frac{1}{\epsilon^C_g} e^C_{ijg,\omega}.$$

Because the scale of utility is not identified from discrete choice models, this is observationally equivalent to the same sub-utility divided through by $1/\epsilon^C_g$ (Train 2009). Doing so yields

$$\ln \varpi^C_{jg} + \epsilon^C_g \ln d^C_{iajg} + e^C_{ijg,\omega}.$$

Given the EV Type I distributional assumption on $e^C_{ijg,\omega}$, choice j is chosen with probability:

$$\frac{\exp(\ln \varpi^C_{jg} + \epsilon^C_g \ln d^C_{iajg})}{\sum_k \exp(\ln \varpi^C_{kg} + \epsilon^C_g \ln d^C_{iakg})} = \frac{\varpi^C_{jg}\left[d^C_{iajg}\right]^{\epsilon^C_g}}{\sum_k \varpi^C_{kg}\left[d^C_{iakg}\right]^{\epsilon^C_g}}.$$

For Equation (8) (the present value sub-utility of a workplace), note that we can right the RHS as

$$\frac{r^*}{\epsilon^W_g} \ln \varpi^W_{jg} + r^* \ln d^W_{iajg} + \frac{r^*}{\epsilon^W_g} \frac{rp}{(1-r)} \ln\left(\Phi^W_{iag}\right) + \frac{r^*}{\epsilon^W_g} e^W_{ijg,\omega}.$$

Because the scale of utility is not identified from discrete choice models, this is observationally equivalent to the same sub-utility divided through by r^*/ϵ^W_g. Doing so yields

$$\ln \varpi^W_{jg} + \epsilon^W_g \ln d^W_{iajg} + \frac{rp}{(1-r)} \ln\left(\Phi^W_{iag}\right) + e^W_{ijg,\omega}.$$

Given the EV Type I distributional assumption on $e^W_{ijg,\omega}$, choice j is chosen with probability:

$$\frac{\exp(\ln \varpi^W_{jg} + \epsilon^W_g \ln d^W_{iajg})}{\sum_k \exp(\ln \varpi^W_{kg} + \epsilon^W_g \ln d^W_{iakg})} = \frac{\varpi^W_{jg}\left[d^W_{iajg}\right]^{\epsilon^W_g}}{\sum_k \varpi^W_{kg}\left[d^W_{iakg}\right]^{\epsilon^W_g}},$$

where the $\frac{rp}{(1-r)} \ln\left(\Phi^W_{iag}\right)$ term drops out because it does not vary with j.

D.1.3 $\mathcal{B}$ recovers β

First, note $\frac{\partial}{\partial \Theta^c_{i(j)gt}} 1[Veh. \in \mathcal{T}_{\omega t}] = \frac{\partial}{\partial \Theta^c_{i(j)gt}} \Pr(a = 1|i, j, y_\omega)$. The law of total probability gives $\Pr(a = 1|i, j, y_\omega) = \sum_g \Pr(a = 1|i, j, g) \Pr(\omega \in g|i, y_\omega)$. Combining these,

$$\begin{aligned}\frac{\partial}{\partial \Theta^c_{i(j)gt}} 1[Veh. \in \mathcal{T}_{\omega t}] &= \sum_g \frac{\partial}{\partial \Theta^c_{i(j)gt}} \Pr(a = 1|i, j, g) \Pr(\omega \in g|i, y_\omega) \\ &= \sum_g \Pr(\omega \in g|i, y_\omega) \mathcal{B}^c_g.\end{aligned}$$

Further, noting that $\frac{\partial}{\partial \Theta^c_{i(j)gt}} \ln \frac{\Pr(a=1|i,j,g)}{\Pr(a=0|i,j,g)} = \beta_g$ and that

$$\frac{\partial}{\partial \Theta^c_{i(j)gt}} \ln \frac{\Pr(a = 1|i, j, g)}{\Pr(a = 0|i, j, g)} = \frac{1}{\Pr(a = 1|i, j, g) \Pr(a = 0|i, j, g)} \frac{\partial \Pr(a = 1|i, j, g)}{\partial \Theta^c_{i(j)gt}},$$

we have

$$\begin{aligned}\mathcal{B}^c_g = \frac{\partial \Pr(a = 1|i, j, g)}{\partial \Theta^c_{i(j)gt}} &= \Pr(a = 1|i, j, g) \Pr(a = 0|i, j, g) \frac{\partial}{\partial \Theta^c_{i(j)gt}} \ln \frac{\Pr(a = 1|i, j, g)}{\Pr(a = 0|i, j, g)} \\ &= \Pr(a = 1|i, j, g) \Pr(a = 0|i, j, g) \beta^c_g.\end{aligned}$$

D.2 Endogenizing Residential Choice

Here we describe one way to endogenize residential location choice that is consistent with Section 5. To do so, we make or maintain several assumptions. First, we assume that households decide residential location prior to vehicle ownership. Second, we assume that a household's potential earnings is not influenced by its residential location choice.[69] Finally, we assume that household spending is not directly influenced by the choice to own a vehicle or not.[70]

Equation 12 describes the utility of car ownership given the initial workplace endowment j and employment status l_ω. While households do not know their future car ownership decision when choosing a residential location, they anticipate optimally choosing vehicle ownership. The expected value of that choice is:

$$W_{ijg}(l_\omega) \equiv \mathbb{E}\left[\max_{a \in \{0,1\}} \mathcal{W}_{iajg,\omega}\left(p_{\text{car}} a, \mathcal{V}^{\text{W}}_{iajg,\omega}, V^{\text{C}}_{iag}, l_\omega\right)\right].$$

Households receive indirect utility $\mathcal{Y}_{ijg,\omega}(l_\omega)$ from residential location i, where

$$\mathcal{Y}_{ijg,\omega}(l_\omega) = f(y_\omega) + \theta_g W_{ijg}(l_\omega) + \gamma_g \ln R_i + \zeta_{ig} + \iota_{ig,\omega},$$

$f(\cdot)$ is a function with $f' > 0$ the marginal utility of income, R_i is the rental price of housing, ζ_{ig} are local amenities, and $\iota_{ig,\omega}$ is an idiosyncratic preference for i. The term $\gamma_g < 0$ scales disutility

69. In our empirical setting, we predict potential earnings based, in part, on residential location due to the availability of data. In alternative settings with richer predictors of income, other approaches that do not condition on residential location may be reasonable.

70. This assumption ensures separability between consumption, housing, and vehicle decisions.

from housing expenses.[71] We assume that $\iota_{ig,\omega}$ is distributed Extreme Value Type I. Under this distributional assumption, the choice probabilities for each residential location i are given by:

$$\pi_{ig} = \frac{\exp\left(\theta_g W_{ijg}(l) + \gamma_g \ln R_i + \zeta_{ig}\right)}{\sum_s \exp\left(\theta_g W_{sjg}(l) + \gamma_g \ln R_s + \zeta_{sg}\right)}.$$

D.2.1 With Observable Rents

Were rents observable, estimation could proceed two steps. First, estimate

$$\pi_{ig} = \frac{\exp\left(\theta_g W_{ijg}(l) + \delta_{ig}\right)}{\sum_s \exp\left(\theta_g W_{sjg}(l) + \delta_{sg}\right)}.$$

where variation across jl identifies θ_g and δ_{ig} are residential location fixed effects. Residential location fixed effects can then be decomposed as

$$\hat{\delta}_{ig} = \gamma_g \ln R_i + \zeta_{ig}.$$

Because ζ_{ig} are likely correlated with rents, some instrument would be useful. One candidate instrument is the presence of a train station. The corresponding first stage would then be:

$$\ln R_i = vD_i + u_i.$$

[[MAKE PANEL VERSIONS]]

D.2.2 Without Observable Rents

In our setting, we lack rental price data that are conformable with the travel surveys. Instead, we estimate a 'reduced form' of the above models:

$$\hat{\delta}_{ig} = \lambda_g D_i + \zeta_{ig},$$

where $\lambda_g = \gamma_g v$. Thus, by examining θ_g, λ_g, and the variation of $W_{ijg}(l)$ in response to transit, we can structurally assess the impact of residential sorting on our results.

71. If, for example, households have Cobb-Douglas preferences over housing and consumption, then $f(y_\omega) \propto \ln(y_\omega)$ and γ_g is the housing expenditure share.

E Additional Effects of Transit on Automobile Adoption and Travel Behavior

E.1 Distance Bins and $\bar{d}$

To determine an appropriate value of $\bar{d}$, we estimate models of transit and vehicle use based on 250m wide bins of distance from nearest rapid transit station to residential tract. If a tract contains a station, the distance is 0m. We estimate these effects on trip-level rapid transit use, trip-level vehicle use, and household vehicle use (this last outcome corresponds to Equation 1). Indexing trips by r, we denote the collection of modes taken for trip r_i that begins at residential tract i as $\mathcal{T}_{rt}$ and estimate:

$$\left.\begin{array}{r} 1[\text{Rapid Transit } \in \mathcal{T}_{rt}] \\ 1[\text{Veh.} \in \mathcal{T}_{rt}] \end{array}\right\} = \sum_k \beta_k 1[\min_{s \in \mathcal{S}_t} d(i,s) \in k] + \delta_i + \varsigma_t + e_{rit} \tag{A1}$$
$$1[\text{Veh.} \in \mathcal{T}_{\omega t}] = \sum_k \beta_k 1[\min_{s \in \mathcal{S}_t} d(i,s) \in k] + \delta_i + \varsigma_t + e_{\omega it}.$$

Here, β_k are the coefficients of interest, corresponding to distance bins k. We estimate binned variants of the baseline design (Equation 1), the PM85 design, and Muni-by-Year FE design (Equation 4), and normalized the effect of distances 5km out or further to 0.

Results are shown in Figure A3. The first panel shows the binned effects on trip-level rapid transit use. Across the three designs, points estimates indicate that receiving a rapid transit station increases rapid transit use by between 18pp and 22pp. This declines steadily to about 1500m away from the station. At that point, rapid transit usage is about 5pp higher than at 5km away from the station. For parsimony, we select $\bar{d} =$ 2kmp-=value for both modes, erring on the side of attenuating our estimates.

The second and third panels show binned estimates of effects on automobile use, at the trip- and household-levels, respectively. There are no systematic patterns in adoption across these two outcomes. Thus we rely on the trip-level adoption results to set $\bar{d}$.

E.2 Returns to Scale for Subway versus BRT

We focus on the similarities and differences in how rapid transit mode (either subway or BRT) impacts travel behavior in our companion piece (Hernández-Cortés, Oliva, and Severen 2024). For the purposes of this project, our primary concern is whether returns to transit density are similar across modes. To this end, we modify Equation 2 to accommodate different effects by mode:

$$1[Veh. \in \mathcal{T}_{\omega t}] = \sum_m \sum_{k=1}^{\bar{k}} \beta_{(k,m)} D_{it,(k)} 1[(k) \text{ is of type } m] + \delta_i + \varsigma_t + e_{\omega it},$$

for $m \in \{\text{Subway}, \text{BRT}\}$.

Cumulative effects, shown in Figure A4, suggest very similar returns to scale regardless of whether the mode is subway or BRT. There is some difference in the effect of the first line; it is significant for subway though insignificant for BRT. However, the cumulative totals for the second through fifth lines tend to be quite similar, with some mild difference in the effect of the fourth line. Regardless, these results support treating BRT and subway similarly when calculating returns to scale.

E.3 Up to 5 Lines of Treatment and Comparisons with Discrete Treatment

We compare the estimates of Equation 2 using the continuous measure of treatment, D_{it} (defined in Equation 3), with a binary measure of proximity. We define the binary measure of proximity as equal to 1 if the station is within 1km of a tract's perimeter and 0 otherwise, splitting the distance of $\bar{d} = 2$km used for D_{it}. Butts (2023) shows that when the magnitude of treatment effects decline in distance, using an indicator for treatment with a too-wide bandwidth attenuates estimates from a ring-type spatial difference-in-differences model.

We test this in our setting, estimating continuous and discrete versions of the baseline and muni-by-year designs of Equation 2. Results, shown in Figure A5, show cumulative effects (rather than marginal effects, which are typically reported in tables). Results reveal attenuation of 30%–50% using the binary measure. We therefore prefer the proximity measure D_{it}, although we cannot use it in all settings.

Additional Results

Figure A1: Plan Maestro 1985 – Subway Construction Plan

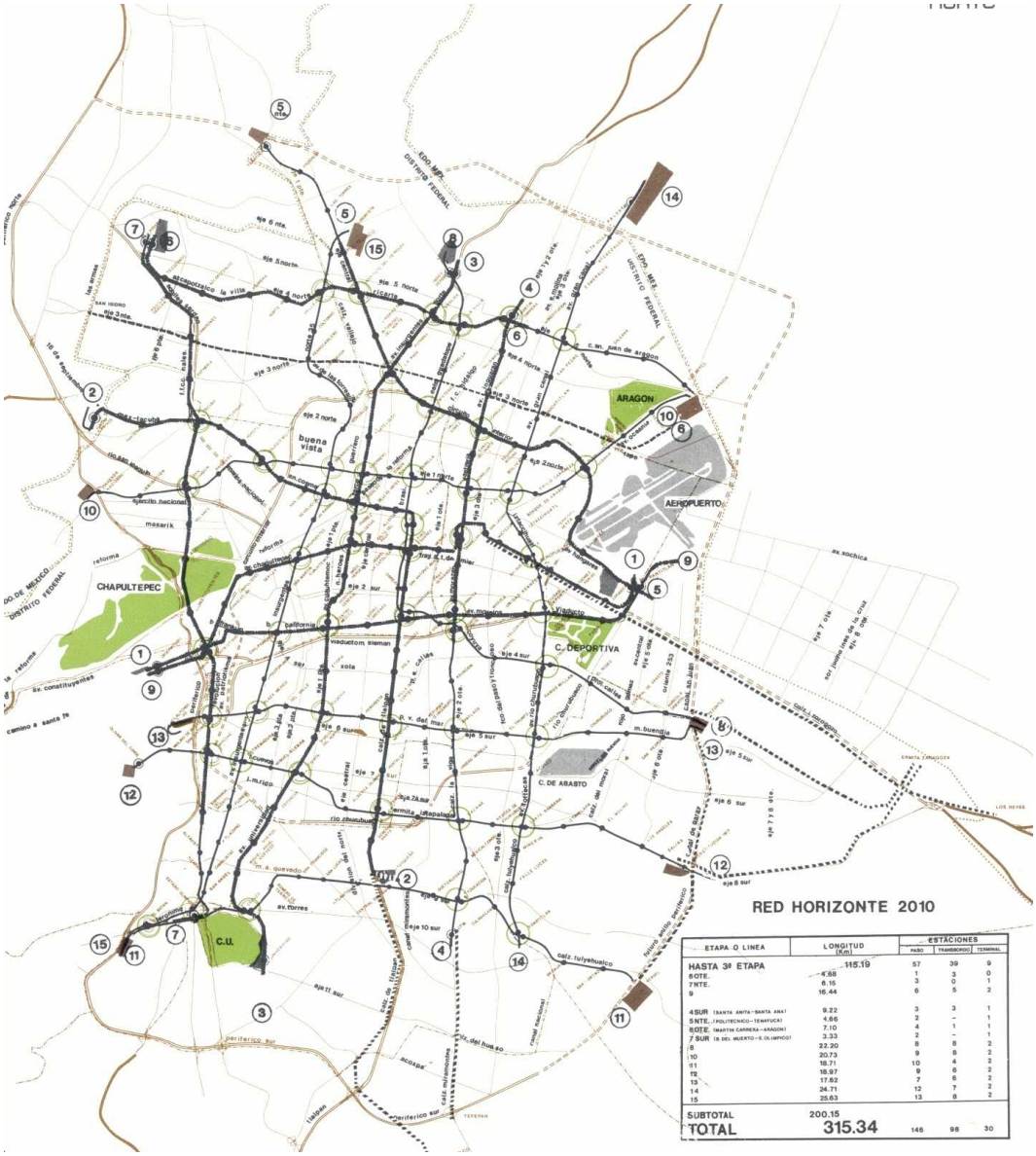


Figure A2: Household-Level Vehicle Usage by Income

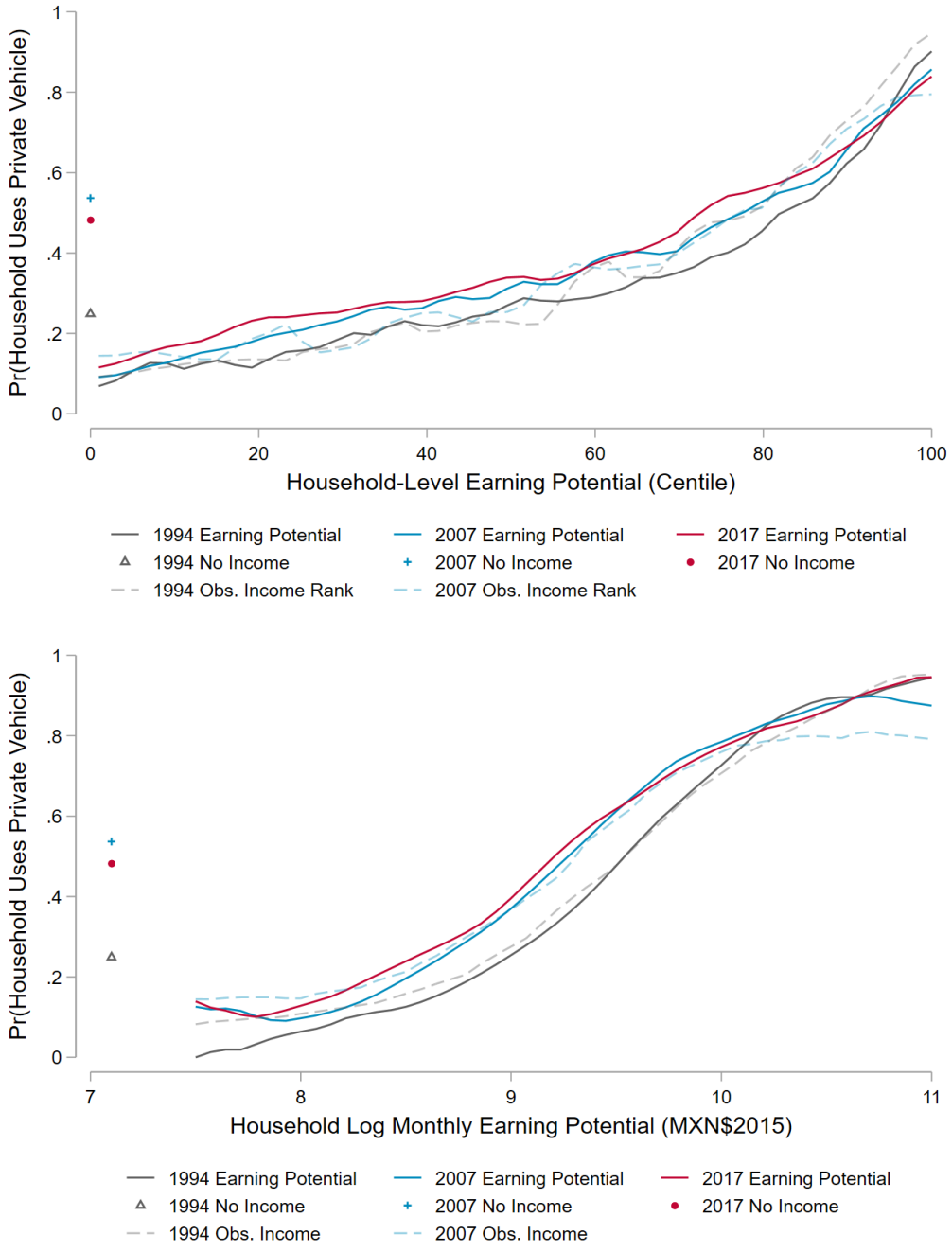


Figure A3: Distance-bin Estimates of Treatment Effects

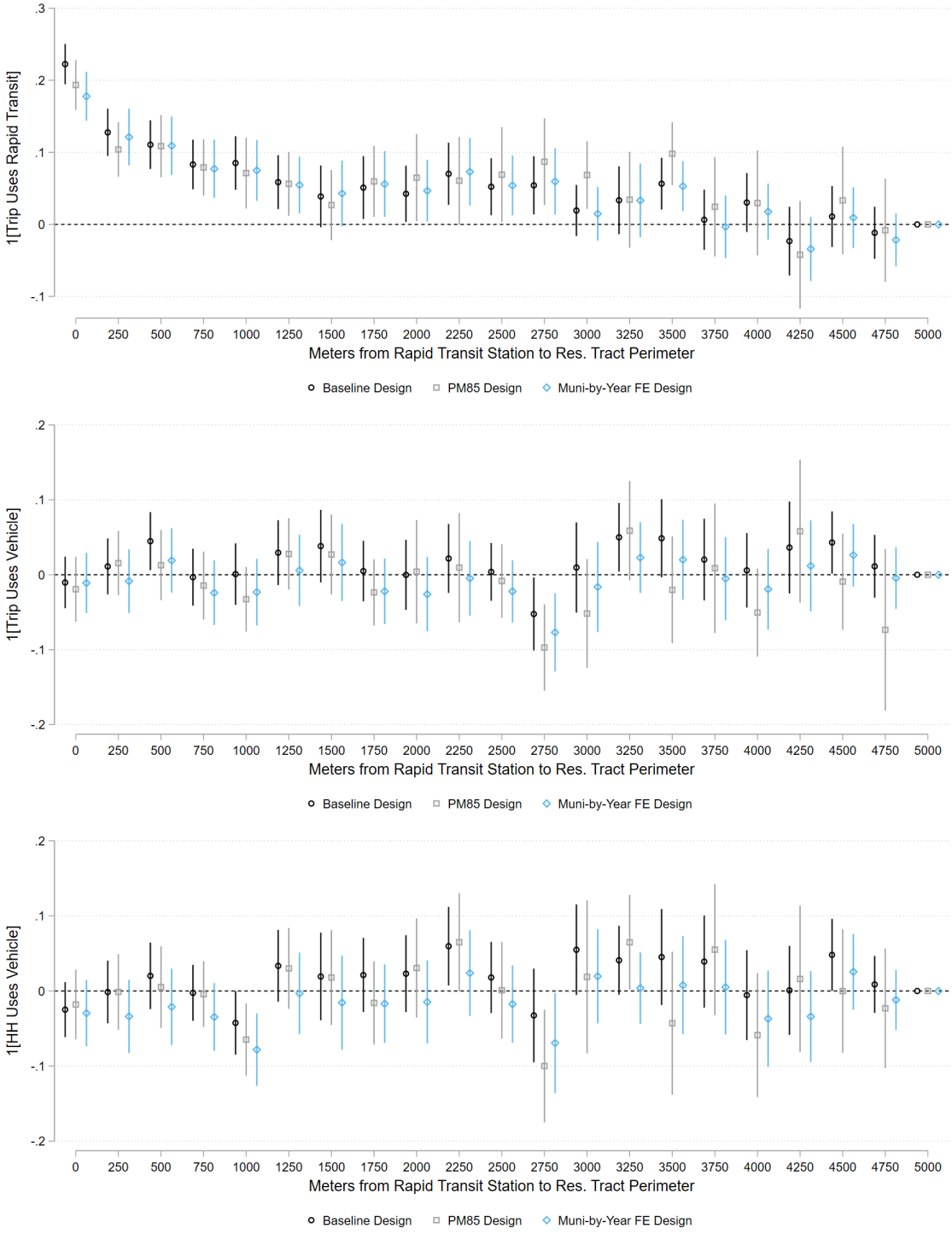
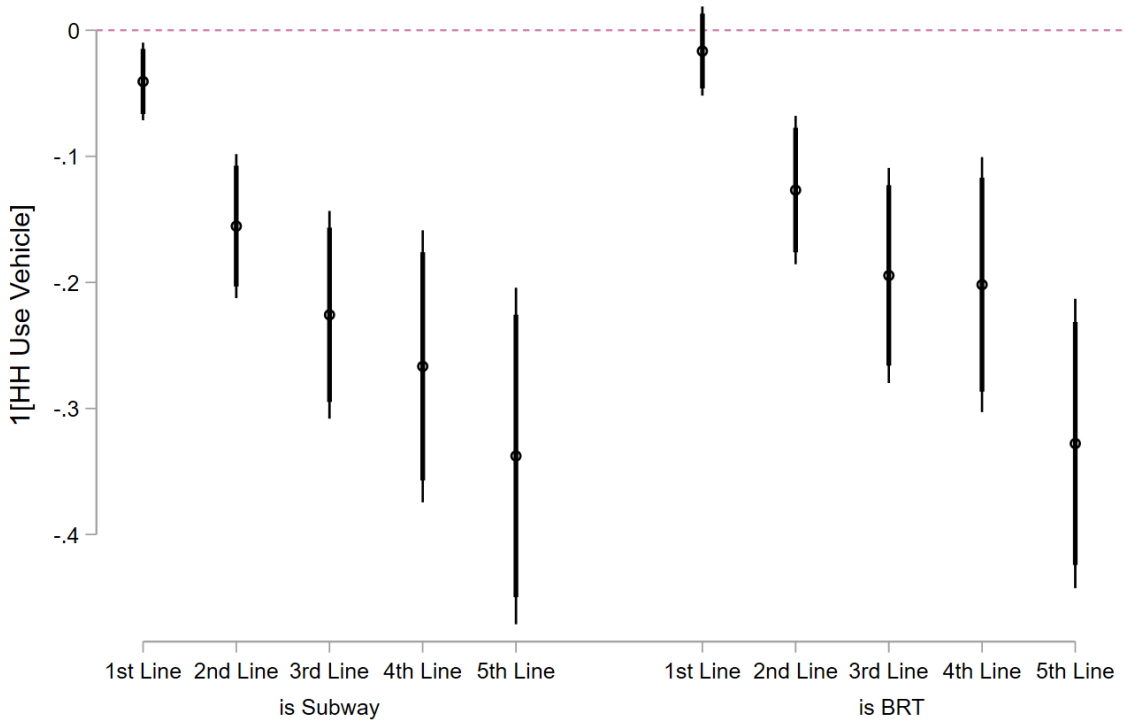


Figure A4: Scale Effects of Transit on Vehicle Use, Comparing Subway and BRT



UPDATE

Figure A5: Returns to Scale and Comparing Continuous and Discrete Treatment

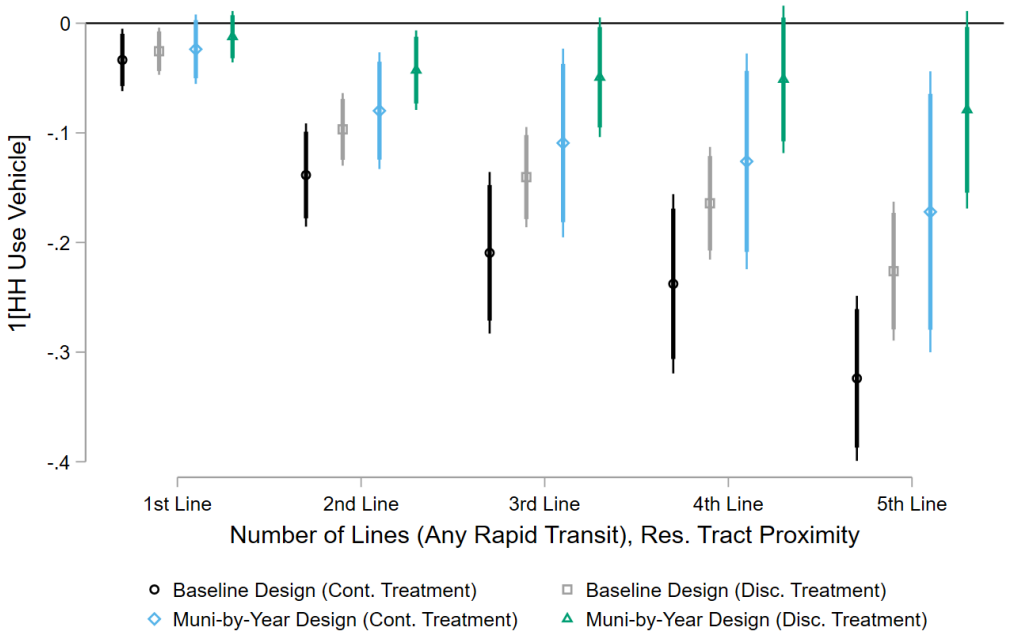


Figure A6: Predicted and Observed Income Densities

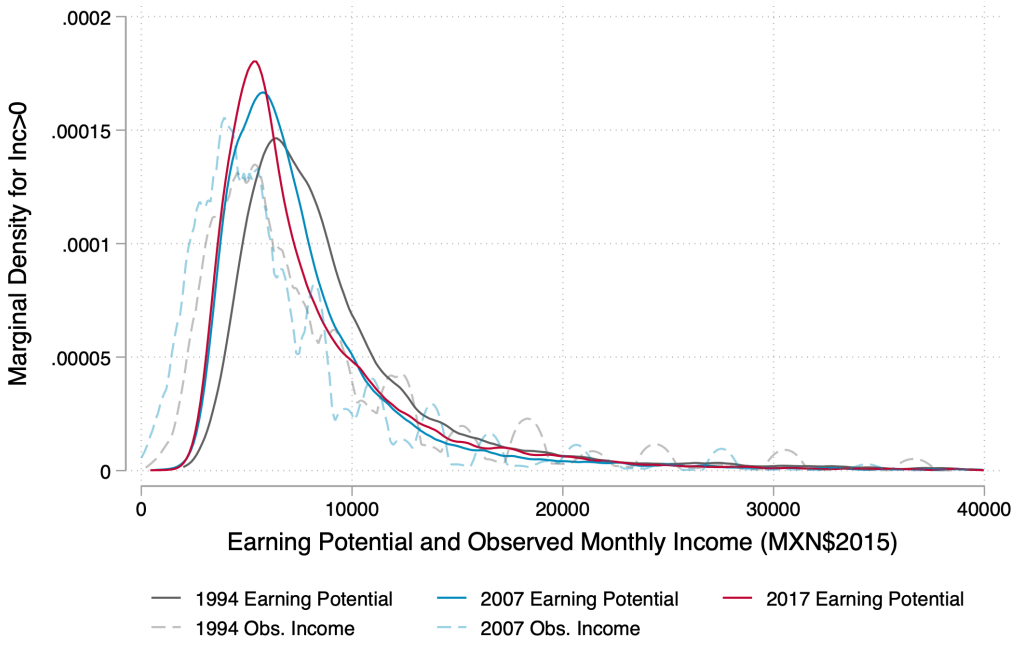


Figure A7: Binscatter of Predicted Income Relative to Observed Income

(a) 1994 Comparison



(b) 2007 Comparison

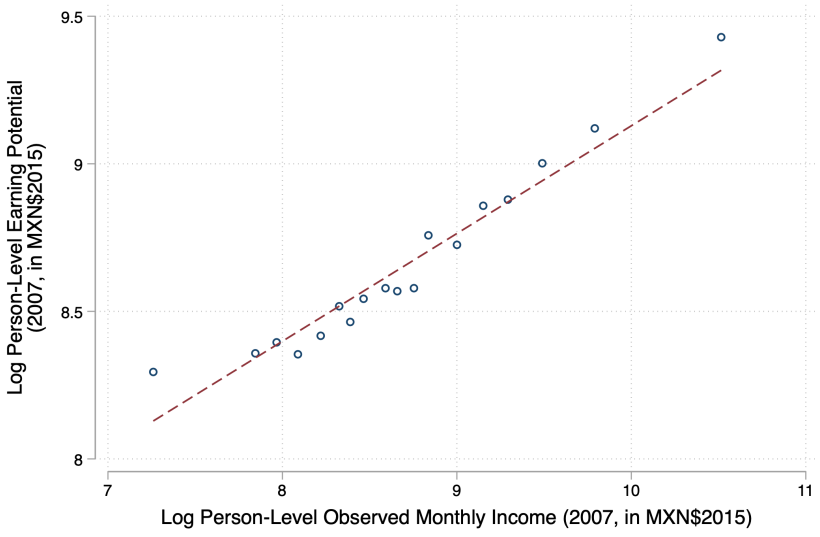


Figure A8: Observed Income Relative to Predicted Income Rank

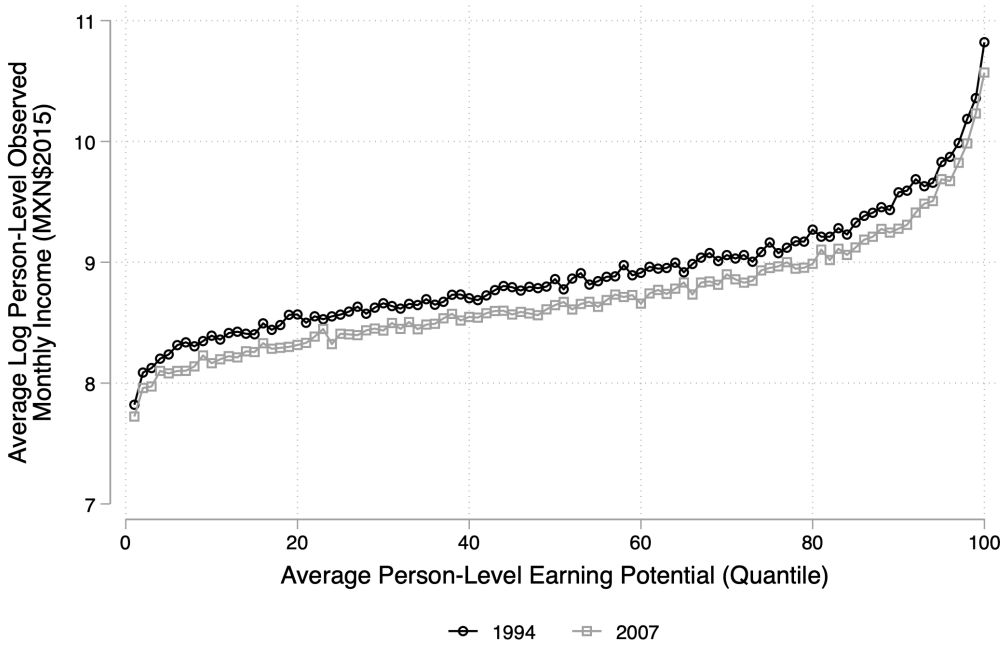


Figure A9: Destination Fixed Effect, 1994 to 2017

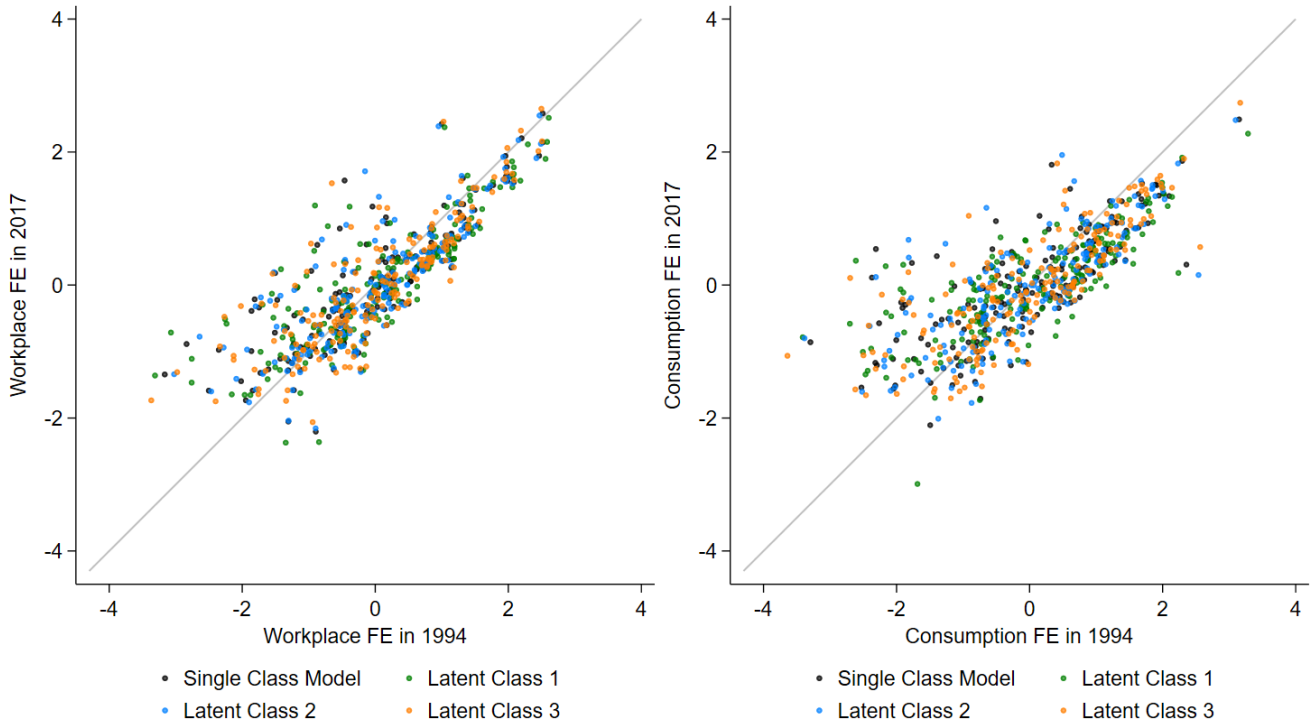
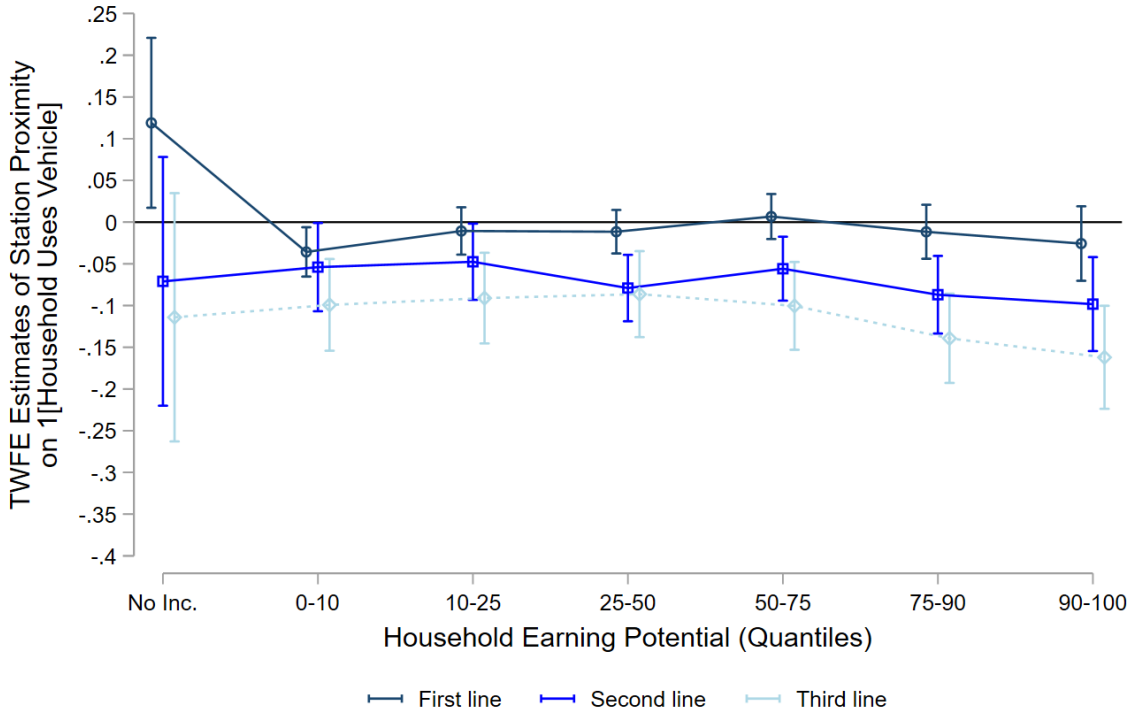


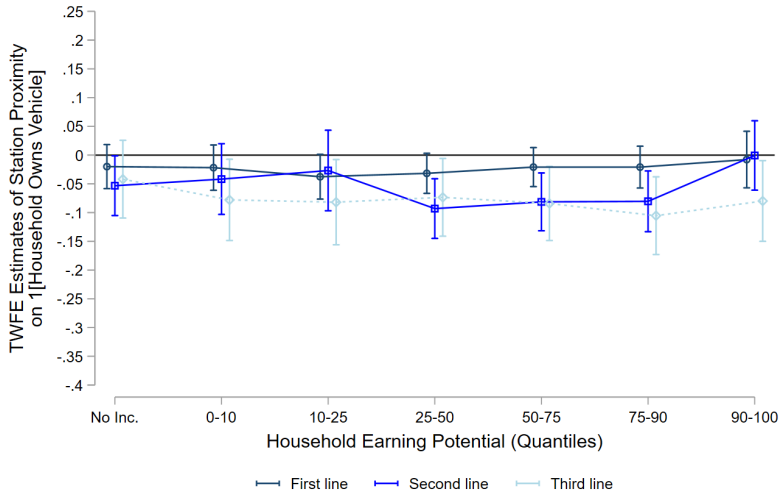
Figure A10: Scale Effects of Transit on Vehicle Use by Potential Income



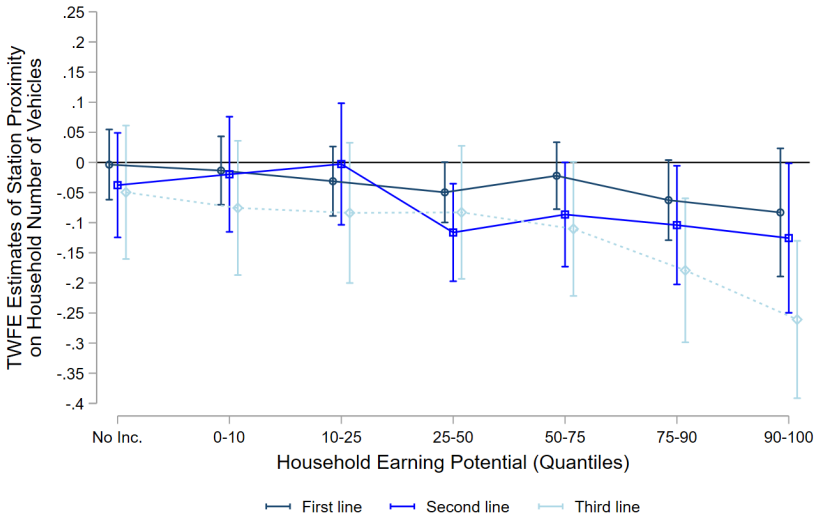
UPDATE

Figure A11: Scale Effects of Transit on Other Vehicle by Potential Income

(a) Household Owns Any Vehicle

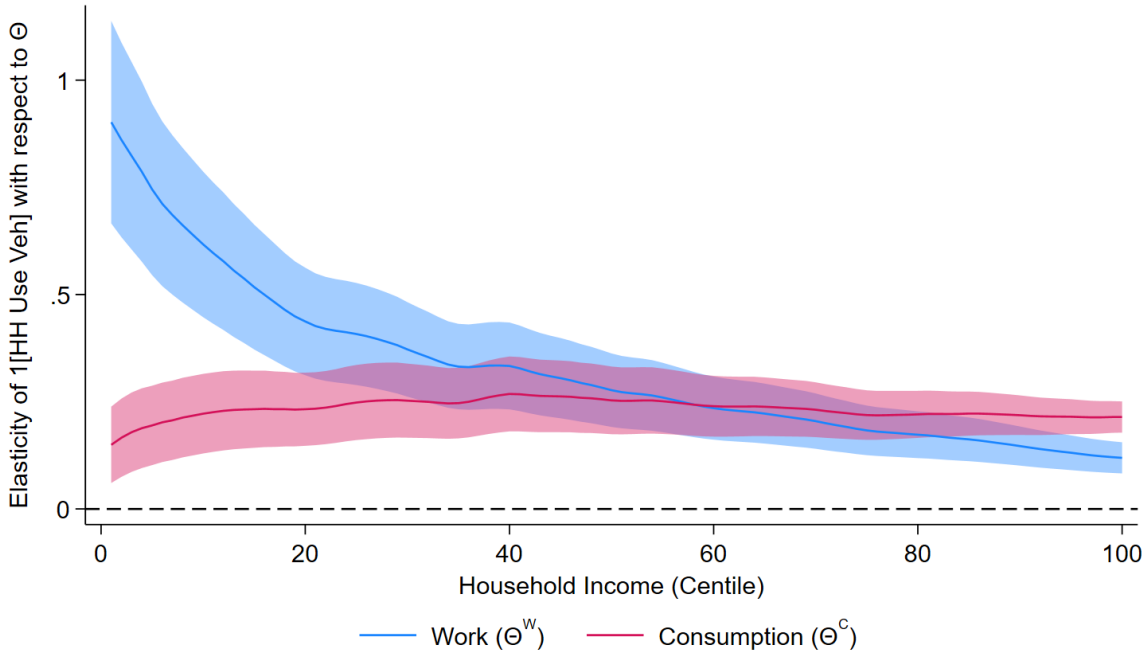


(b) Number of Vehicles in Household



UPDATE

Figure A12: Elasticity of Vehicle Adoption with respect to Θ



UPDATE

Table A1: Timeline

Time	Event
<1994	Metro Lines 1–7, 9, and A
5/1994	**EOD Survey 1**
7/1994	Metro Line 8 Opens
12/1999	Metro Line B Opens (first 13 stations)
11/2000	Metro Line B Expands (final 8 stations)
6/2005	Metrobús Line 1 Opens (first 37 stations)
5/2007	**EOD Survey 2**
3/2008	Metrobús Line 1 Expands (final 9 stations)
12/2008	Metrobús Line 2 Opens
2/2011	Metrobús Line 3 Opens
4/2012	Metrobús Line 4 Opens
10/2012	Metro Line 12 Opens
11/2013	Metrobús Line 5 Opens
1/2016	Metrobús Line 6 Opens
2/2017	**EOD Survey 3**

Table A2: Distribution of Treatment Across Tracts

No. of Lines in 1994	No. of Lines in 2017				Totals in 1994
	0	1	2	3+	
0	55.93%	13.68%	2.66%	0.04%	72.32%
1	-	5.85%	5.29%	2.87%	14.00%
2	-	-	2.10%	7.26%	9.36%
3+	-	-	-	4.32%	4.32%
Totals by 2017	55.93%	19.53%	10.05%	14.49%	

Table A3: Distribution of Treatment Across Tracts

No. of Lines in 1994	No. of Lines gained 1994–2017						Totals in 1994
	0	1	2	3	4	5	
0	55.93%	13.68%	2.66%	0.04%	0.00%	0.00%	72.32%
1	5.85%	5.29%	2.30%	0.44%	0.08%	0.04%	14.00%
2	2.10%	2.30%	3.35%	1.21%	0.40%	0.00%	9.36%
3	0.52%	1.37%	1.41%	1.01%	0.00%	0.00%	4.32%
Totals gains 1994–2017	64.41%	22.64%	9.73%	2.70%	0.48%	0.04%	

Table A4: Scale Effects in Other Vehicle Outcomes (Baseline and Muni Designs)

	1[HH Owns Vehicle]				Number of Vehicles in HH			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1(+) Line(s)	-0.047*	-0.045*	-0.037+	-0.037+	-0.065*	-0.063*	-0.049	-0.050
	(0.019)	(0.019)	(0.022)	(0.022)	(0.031)	(0.031)	(0.035)	(0.035)
2 Lines		-0.035		-0.010		-0.021		0.022
		(0.024)		(0.028)		(0.040)		(0.046)
3+ Lines		-0.023		-0.044		-0.067		-0.116+
		(0.029)		(0.037)		(0.052)		(0.063)
Cumulative Eff.		-0.104**		-0.091+		-0.151*		-0.143
		(0.040)		(0.053)		(0.070)		(0.088)
Pr(CRS of 2nd Line)		[0.764]		[0.430]		[0.415]		[0.203]
Pr(CRS of 3rd+ Line)		[0.626]		[0.607]		[0.677]		[0.137]
Years	Ex. '94	Ex. '94	Ex. '94	Ex. '94	Ex. '94	Ex. '94	Ex. '94	Ex. '94
Design	Baseline	Baseline	Muni	Muni	Baseline	Baseline	Muni	Muni
Tract FEs	Y	Y	Y	Y	Y	Y	Y	Y
Year FEs	Y	-	Y	-	Y	-	Y	-
Muni×Year FEs	-	Y	-	Y	-	Y	-	Y

81836 observations. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table A5: Scale Effects in Other Vehicle Outcomes (PM85 and Scaffolded Designs)

	1[HH Owns Vehicle]			Num. of Veh. in HH		
	(1)	(2)	(3)	(4)	(5)	(6)
1st (1+) Line | 0 Lines Present	-0.043+	-0.053**	-0.049*	-0.051	-0.079**	-0.078*
	(0.023)	(0.020)	(0.024)	(0.038)	(0.033)	(0.038)
2nd Line | 1 Line Present		0.007	0.036		0.087+	0.105
		(0.033)	(0.041)		(0.053)	(0.068)
3rd Line | 2 Lines Present		-0.200***	-0.244***		-0.245**	-0.325**
		(0.050)	(0.061)		(0.086)	(0.102)
Cumulative Eff.		-0.246***	-0.257***		-0.236*	-0.298*
		(0.063)	(0.078)		(0.106)	(0.129)
Pr(CRS of 2nd Line)		[0.123]	[0.074]		[0.007]	[0.020]
Pr(CRS of 3rd Line)		[0.001]	[0.000]		[0.006]	[0.002]
Design	PM85	Scaffold	Scaffold	PM85	Scaffold	Scaffold
Tract FEs	Y	-	-	Y	-	-
Tract×Scaffold FEs	-	Y	Y	-	Y	Y
Year FEs	Y	-	-	Y	-	-
Year×Scaffold FEs	-	Y	-	-	Y	-
Muni×Year×Scaffold FEs	-	-	Y	-	-	Y
N	34048	81735	81735	34048	81735	81735

UPDATE

Table A6: Long Differences (Omitting 2007)

	1[Any HH Vehicle Use]			
	(1)	(2)	(3)	(4)
1(+) Line(s)	-0.051**	-0.036+	-0.039+	-0.031
	(0.018)	(0.020)	(0.021)	(0.022)
2 Lines		-0.136***		-0.079*
		(0.035)		(0.038)
3+ Lines		-0.120***		-0.019
		(0.033)		(0.042)
Cumulative Eff.		-0.292***		-0.129*
		(0.038)		(0.052)
Pr(CRS of 2nd Line)		[0.032]		[0.347]
Pr(CRS of 3rd+ Line)		[0.423]		[0.455]
Design	Baseline	Baseline	Muni	Muni
Tract FEs	Y	Y	Y	Y
Year FEs	Y	-	Y	-
Muni×Year FEs	-	Y	-	Y

57710 observations. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table A7: Effects for Households with Younger Householders Versus Older Householders

	1[Any HH Vehicle Use]			
	(1)	(2)	(3)	(4)
1(+) Line(s)	-0.065**	-0.055**	-0.019	-0.013
	(0.020)	(0.021)	(0.019)	(0.019)
2 Lines		-0.134***		-0.090**
		(0.034)		(0.029)
3+ Lines		-0.137***		-0.068+
		(0.039)		(0.036)
Cumulative Eff.		-0.325***		-0.171***
		(0.047)		(0.043)
Pr(CRS of 2nd Line)		[0.070]		[0.038]
Pr(CRS of 3rd+ Line)		[0.368]		[0.697]
Sample	HoH≤40	HoH≤40	HoH>40	HoH>40
Design	Baseline	Baseline	Baseline	Baseline
Tract FEs	Y	Y	Y	Y
Year FEs	Y	Y	Y	Y
N	43458	43458	47783	47783

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table A8: Mini-Experiments

	1[Any HH Vehicle Use]			1[HH Owns Vehicle]			Number of Vehicles in HH		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Experiment									
Access	-0.004	0.018	0.012	-0.003	0.023	0.007	-0.016	0.026	-0.004
	(0.014)	(0.018)	(0.015)	(0.016)	(0.019)	(0.018)	(0.026)	(0.029)	(0.028)
Connectivity	-0.141***	-0.128***	-0.093***	-0.066**	-0.079*	-0.053*	-0.095**	-0.110*	-0.066+
	(0.022)	(0.029)	(0.022)	(0.023)	(0.031)	(0.025)	(0.036)	(0.046)	(0.038)
Redundancy	-0.058+	0.042	0.005	-0.060+	0.042	-0.063+	-0.064	0.086	-0.077
	(0.032)	(0.061)	(0.034)	(0.033)	(0.060)	(0.036)	(0.054)	(0.089)	(0.061)
Other Station	-0.024*	-0.013	-0.016	-0.038*	-0.034+	-0.028+	-0.057*	-0.052+	-0.041
	(0.011)	(0.013)	(0.012)	(0.015)	(0.018)	(0.017)	(0.024)	(0.029)	(0.025)
Mean Outcome	0.357	0.364	0.357	0.446	0.449	0.446	0.591	0.577	0.591
Years	All	All	All	'07-'17	'07-'17	'07-'17	'07-'17	'07-'17	'07-'17
Design	Baseline	PM85	Muni	Baseline	PM85	Muni	Baseline	PM85	Muni
Tract FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FEs	Y	Y	-	Y	Y	-	Y	Y	-
Muni×Year FEs	-	-	Y	-	-	Y	-	-	Y
N	91256	39114	91256	81836	34048	81836	81836	34048	81836

Experiments are not exclusive. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table A9: Income Summary Statistics

		1994	2007	2017
Observed Income	mean	9594.7	8009.1	-
	sd	10909.7	11365.6	-
Earning Potential	mean	9666.8	8011.0	7924.9
	sd	7557.1	6162.3	5641.7
Observations		28631	47117	50174

All figures in MXN$2015.

Table A10: Testing for Sorting on Observables

	ln(Income) (actual) (1)	ln(Income) (pred.) (2)	1[Working] (3)	Age (4)	1[Female] (5)	1[Student] (6)	1[≥HS] (7)	1[≥College] (8)
Panel A. Baseline Design								
$D_{\text{Any Rapid},ct}$	0.027	-0.038+	-0.003	1.087***	0.013	-0.007	-0.018	-0.027
	(0.037)	(0.021)	(0.011)	(0.302)	(0.008)	(0.009)	(0.019)	(0.017)
N	67748	123799	184686	184709	184709	184686	106265	106265
Panel B. PM85 Design								
$D_{\text{Any Rapid},ct}$	0.080+	-0.011	-0.023+	0.789*	0.020*	0.003	0.002	-0.016
	(0.041)	(0.025)	(0.014)	(0.358)	(0.010)	(0.011)	(0.022)	(0.020)
N	31208	54432	80181	80193	80193	80181	45354	45354
Panel C. Muni Design								
$D_{\text{Any Rapid},ct}$	0.005	-0.038	0.013	0.892**	0.003	-0.014	-0.008	-0.022
	(0.040)	(0.024)	(0.013)	(0.338)	(0.009)	(0.010)	(0.021)	(0.019)
N	67748	123798	184686	184709	184709	184686	106265	106265
Sample	Workers	Workers	All	All	All	All	≥25yo	≥25yo
Years	'94 & '07	All	All	All	All	All	'07 & '17	'07 & '17

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 A11: Testing for Sorting on Observables with work location FE

	ln(Income) (actual) (1)	ln(Income) (pred.) (2)	1[Working] (3)	Age (4)	1[Female] (5)	1[Student] (6)	1[≥HS] (7)	1[≥College] (8)
Panel A. Baseline Design								
$D^{Station}_ct$	0.040	-0.034+	-0.002	1.117***	0.016+	-0.008	-0.019	-0.030+
	(0.035)	(0.020)	(0.012)	(0.311)	(0.009)	(0.009)	(0.019)	(0.017)
N	64824	164754	167379	167398	167398	167379	94032	94032
Panel C. Muni Design								
$D^{Station}_ct$	0.013	-0.033	0.008	0.931**	0.009	-0.015	-0.007	-0.024
	(0.037)	(0.023)	(0.013)	(0.352)	(0.010)	(0.010)	(0.021)	(0.019)
N	64824	164754	167379	167398	167398	167379	94032	94032
Sample	Workers	Workers	All	All	All	All	≥25yo	≥25yo
Years	'94 & '07	All	All	All	All	All	'07 & '17	'07 & '17

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 A12: Trip purpose and destinations

	1994 %	2007 %	2017 %	yr17 %
Purpose: work	0.546	0.510	0.534	0.593
Purpose: shopping	0.083	0.084	0.087	0.078
Purpose: school	0.150	0.188	0.127	0.135
Purpose: social	0.044	0.045	0.040	0.046
Purpose: accompany	0.087	0.082	0.096	0.084
Observations	199,764	48,382	79,347	72,035
Panel B: Destination for work purpose				
Office	0.365	0.356	0.331	0.402
Factory	0.180	0.233	0.158	0.155
Shops	0.175	0.161	0.144	0.214
School	0.060	0.055	0.071	0.054
Sport facility	0.003	0.003	0.004	0.002
Medical facility	0.041	0.042	0.044	0.039
Dining	0.019	0.022	0.026	0.011
Other	0.157	0.130	0.220	0.123
Observations	112935	28070	42232	42633
Panel C: Destination for shopping purpose				
Office	0.002	0.000	0.002	0.003
Factory	0.007	0.003	0.006	0.011
Shops	0.962	0.983	0.939	0.965
School	0.001	0.000	0.001	0.001
Sport facility	0.000	0.000	0.000	0.000
Medical facility	0.001	0.000	0.000	0.001
Dining	0.001	0.000	0.001	0.001
Other	0.027	0.014	0.050	0.017
Observations	16103	3573	6913	5617
Panel D: Destination for school purpose				
Office	0.002	0.000	0.002	0.005
Factory	0.001	0.000	0.001	0.001
Shops	0.001	0.000	0.001	0.003
School	0.986	0.998	0.979	0.975
Sport facility	0.000	0.000	0.001	0.000
Medical facility	0.005	0.002	0.007	0.006
Dining	0.000	0.000	0.000	0.000
Other	0.005	0.000	0.008	0.010
Observations	27868	7800	10190	9878
Panel E: Destination for social purpose				
Office	0.006	0.006	0.009	0.004
Factory	0.002	0.001	0.004	0.002
Shops	0.057	0.017	0.064	0.090
School	0.014	0.007	0.019	0.016
Sport facility	0.090	0.084	0.111	0.078
Medical facility	0.025	0.007	0.026	0.040
Dining	0.037	0.023	0.043	0.046
Other	0.769	0.854	0.724	0.724
Observations	8163	1760	3134	3269
Panel F: Destination for accompany purpose				
Office	0.026	0.031	0.024	0.024
Factory	0.006	0.007	0.006	0.004
Shops	0.016	0.005	0.014	0.028
School	0.797	0.830	0.780	0.784
Sport facility	0.005	0.004	0.009	0.001
Medical facility	0.065	0.057	0.065	0.073
Dining	0.002	0.002	0.002	0.002
Other	0.084	0.066	0.101	0.084
Observations	16877	3409	7497	5971

Table A13: Neighborhood Quintile-Quintile Income Transition Matrix

Neighborhood Earning Potential (Quintile), 1994	Neighborhood Earning Potential (Quintile), 2017				
	1	2	3	4	5
1	7.7%	6.1%	3.7%	2.0%	0.6%
2	5.4%	5.4%	5.3%	3.1%	1.0%
3	3.8%	4.5%	5.6%	3.9%	1.9%
4	2.6%	2.9%	3.9%	7.2%	3.3%
5	0.5%	1.2%	1.6%	3.8%	13.1%

Table A14: Head of Household vs. not Head of Household

	1[Any Rapid Tran.]		1[Any Bus/Coll.]		1[Any Transit]		1[Any Vehicle]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$D_{\text{Any Rapid},ct}$	0.158***	0.196***	-0.134***	-0.160***	0.010	0.007	-0.029+	-0.017
	(0.015)	(0.014)	(0.016)	(0.015)	(0.015)	(0.013)	(0.016)	(0.012)
Mean Outcome	0.253	0.284	0.607	0.705	0.649	0.759	0.321	0.222
Tract FEs	Y	Y	Y	Y	Y	Y	Y	Y
Year FEs	Y	Y	Y	Y	Y	Y	Y	Y
Heads of Household	Yes	No	Yes	No	Yes	No	Yes	No
N	79846	104863	79846	104863	79846	104863	79846	104863

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table A16: Standard Deviation of Destination Value

	Single Class	Latent Class		
		Low	Mid	High
Panel A. Workplace Fixed Effects				
1994	1.089	1.151	1.069	1.103
2007	.883	.846	.903	.908
2017	.908	.903	.909	.928
Panel B. Consumption Destination Fixed Effects				
1994	1.147	1.194	1.132	1.199
2007	.888	.971	.872	.923
2017	.807	.743	.816	.869

Standard deviation of destination fixed effects.

Table A15: Model Estimates – Mode Choice (predicting no income)

	Single Class		Latent Class					
	Work	Cons.	Work			Consumption		
			Low	Mid	High	Low	Mid	High
	(1a)	(1b)	(2a)	(2b)	(2c)	(2d)	(2e)	(2f)
A. Travel Time Disutility								
κ	-0.0117	-0.0095	-0.0133	-0.0121	-0.0220	-0.0054	-0.0098	-0.0176
$\kappa_{Taxi,94}$	-0.2899	-0.3119	-0.4288	-0.2897	-0.2204	-0.4015	-0.3817	-0.2223
$\kappa_{Car,94}$	-0.0617	-0.0928	-0.1080	-0.0576	-0.0871	-0.0583	-0.0942	-0.1022
$\kappa_{Taxi,07}$	-0.2033	-0.1876	-0.2913	-0.1981	-0.2088	-0.1867	-0.1926	-0.2155
$\kappa_{Car,07}$	-0.0621	-0.0610	-0.0702	-0.0665	-0.0985	-0.0324	-0.0517	-0.0850
$\kappa_{Taxi,17}$	-0.1964	-0.1612	-0.2275	-0.2158	-0.1806	-0.1172	-0.1948	-0.1670
$\kappa_{Car,17}$	-0.0594	-0.0561	-0.0873	-0.0591	-0.0839	-0.0014	-0.0543	-0.0860
B. Nesting Parameters								
λ_{Public}	0.2507	0.1773	0.5364	0.1857	0.3138	0.1626	0.1513	0.1611
$\lambda_{RapidTrans}$	0.1221	0.0952	0.1392	0.1120	0.1889	0.0719	0.0659	0.1645
$\lambda_{LastMode}$	0.1474	0.1130	0.1495	0.1467	0.2807	0.0664	0.1102	0.1903
$\lambda_{Private}$	1.2823	1.1644	1.2387	1.2107	1.2526	1.0153	1.1857	1.0856
C. Mode Utility Shifters								
b_{Sub}	0	0	0	0	0	0	0	0
$b_{Sub\times BRT}$	-0.3255	-0.2939	-0.4983	-0.3530	-0.5700	-0.1512	-0.2987	-0.4252
$b_{Sub\times Inf}$	0.0050	-0.0430	0.3016	-0.0443	-0.2427	0.0487	-0.0462	-0.2331
$b_{BRT\times Sub}$	-0.2387	-0.1978	-0.1862	-0.1946	-0.6306	-0.1246	-0.1227	-0.5660
b_{BRT}	0.0520	0.0139	0.1088	0.0085	0.1864	0.0036	0.0280	0.0495
$b_{BRT\times Inf}$	-0.2299	-0.2422	-0.1957	-0.2495	-0.4193	-0.1514	-0.1921	-0.4027
$b_{Inf\times Sub}$	0.1751	0.0798	0.5172	0.1559	-0.1520	0.1416	0.0904	-0.1702
$b_{Inf\times BRT}$	-0.0927	-0.1074	0.0296	-0.1053	-0.4347	-0.0451	-0.0901	-0.3249
b_{Inf}	0.2523	0.1865	0.6319	0.2327	-0.0530	0.2051	0.1964	0.0071
b_{Taxi}	-0.9912	-0.6901	-0.3029	-1.1546	-0.9038	-0.5149	-0.7445	-0.4624
b_{Car}	1.4225	0.9522	1.5568	1.0462	2.0075	0.6796	0.3969	1.6381
$ME1$	-2.2631	-1.8154	-2.9557	-2.2470	-4.2469	-1.0402	-1.8646	-3.1475
$ME2$	0.0238	0.0645	0.1980	-0.1427	0.2610	0.2154	-0.0618	0.0881
D. Class Assignment Parameters								
μ_1					10.6563			
μ_2					4.3558			
μ_f					2.5302			
σ_f					0.1600			
μ_{mi94}					-14.3872			
μ_{mi07}					17.0843			
μ_{mi17}					15.2512			
Class Probabilities			Pr(Low)	Pr(Mid)	Pr(Hi.)			
$F(y) = 0.005$			0.9123	0.0865	0.0012			
$F(y) = 0.1$			0.6056	0.3862	0.0083			
$F(y) = 0.25$			0.3510	0.6258	0.0232			
$F(y) = 0.5$			0.1302	0.7908	0.0790			
$F(y) = 0.75$			0.0345	0.7015	0.2640			
$F(y) = 0.9$			0.0085	0.3925	0.5990			
$F(y) = 0.995$			0.0002	0.0187	0.9811			
LL (prior)	update	update			1.9674E+07			
LL (final, no penal.)	update	update			1.9573E+07			
Penalization	-				0.01			

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 A17: Spatial Autocorrelation of Workplace and Consumption Destination Value

	Single Class	Latent Class		
		Low	Mid	High
Panel A. Workplace Fixed Effects				
1994	0.616	0.639	0.616	0.613
2007	0.634	0.600	0.643	0.633
2017	0.576	0.565	0.580	0.573
Panel B. Consumption Destination Fixed Effects				
1994	0.532	0.520	0.524	0.547
2007	0.447	0.442	0.449	0.483
2017	0.461	0.333	0.465	0.503

Spatial autocorrelation of destination FEs calculated using Moran's I.