Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification*

Christopher Severen†
Federal Reserve Bank of Philadelphia

October 2019

Abstract

I estimate that a large investment in rail transit in Los Angeles between 1990 and 2000 has a positive effect on commuting using panel data on bilateral flows. Commuting between location pairs that both contain stations increases by 15%. I use a spatial general equilibrium model to isolate non-commuting effects of transit and measure welfare. Local innovations interacted with intraurban geography identify key model parameters; estimates suggest inelastic labor mobility and housing supply. Metro Rail increases welfare $146 million annually by 2000, less than operational subsidies and annualized capital costs. More recent data show some additional commuting growth.

Keywords: subway, commuting, gravity, economic geography, local labor supply

JEL Codes: J61, L91, R13, R31, R40

*Many have provided valuable advice during the development of this paper; I would especially like to thank Elliot Anenberg, Valerie Bostwick, Jeff Brinkman, Olivier Deschênes, Jonathan Dingel, Gilles Duranton, Jessie Handbury, Stephan Heblich, Kilian Heilmann, Nicolai Kuminoff, Jeff Lin, Kyle Meng, Yuhei Miyauchi, Paulina Oliva, and Wei You, as well as several anonymous referees. First draft was December 10, 2015.

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

†Address: Research Department, Federal Reserve Bank of Philadelphia, Ten Independence Mall, Philadelphia, PA 19106. Email: chris.severen@gmail.com or Chris.Severen@phil.frb.org.

1 Introduction

High commuting costs limit consumer choice and mobility within cities. Governments invest large sums in rail transit infrastructure to mitigate the costs of distance and congestion. What are the benefits of these investments, and do these benefits come from changes in commuting behavior or other margins?

I study the effects of Los Angeles Metro Rail on commuting, non-commuting margins, and welfare. I first assemble data on census tract-to-census tract commuting flows and travel times in 1990 and 2000 and develop an estimation framework that uses bilateral and panel aspects of the data in conjunction with shocks to routing to recover the commuting effect of transit. To quantify other margins of effects and quantify welfare impacts, I then turn to a quantitative spatial general equilibrium model and develop a new strategy to identify its key parameters. A gravity equation unites analysis of the commuting effect and the structural model. This equation includes includes origin-year and destination-year fixed effects, aiding identification of the model and allaying selection concerns due to non-random placement of transportation infrastructure.[1] Instead of comparing single locations, identification of the commuting effect hinges on selecting *pairs* of locations that satisfy treatment ignorability. In practice, this means comparing changes in flows between pairs of locations that both receive treatment to changes in flows between pairs of locations in which just one, or neither, receives treatment.

I exploit unanticipated shocks to route construction and historical maps of streetcar and proposed subway lines to find plausible control pairs and estimate the causal impact of transit on commuting flows. The commuting effect of LA Metro Rail is substantial: Commuting increases by 11%-15% between connected tract pairs that both contain stations by 2000. Slightly more distant pairs show increases of 9%-13%, and there is no effect on tracts further away. I also find some evidence of marginal congestion improvements in response to LA Metro Rail network by 2000. The methods used contrast those common in the urban economic geography literature, which typically uses a single cross section of flow data to infer changes in commuting from changes in computer-generated travel time (e.g., Ahlfeldt et al. 2015; Allen, Arkolakis, and Li 2015; Monte, Redding, and Rossi-Hansberg 2018).

I incorporate observed commuting flows into a quantitative spatial general equilibrium model of Los Angeles to distinguish commuting from non-commuting effects of transit. Conditional on model parameters and panel bilateral commuting data, panel data on local housing and labor prices map to time-varying, tract-specific primitives; these represent non-commuting fundamentals (such as productivity and amenities). I estimate the effect of transit on changes in these fundamentals with a differences-in-differences strategy, which controls for potentially confounding time-invariant factors (such as proximity to natural amenities). This strategy explicitly tests whether transit infrastructure alters city structure through non-commuting channels, bridging

1. See Redding and Turner (2015) for a review of this challenge and common solutions.

the hedonic approach to valuing transportation infrastructure (e.g., Baum-Snow and Kahn 2000; McMillen and McDonald 2004) with that taken by quantitative urban models (e.g., Monte, Redding, and Rossi-Hansberg 2018; Tsivanidis 2018). The commuting effect dominates; impacts from non-commuting channels appear minimal. This is perhaps surprising, as non-commuting transit-related amenities and productivity effects are often discussed.[2]

The model also quantifies the effects of transit infrastructure in terms of welfare, and can account for general equilibrium effects.[3] I develop a new identification strategy for key model parameters. Foremost is the local (extensive-margin) elasticity of labor supply to a tract. This elasticity governs how responsive agents are to changes in prices, amenities, and commuting costs, and is essential to translate treatment effects to utility. To identify this parameter, most urban economic geography models require that, conditional on residential geography and travel time, either there are no factors that shift labor supply or that any such factors are orthogonal to workplace wage. Such assumptions are often necessary because workplace wage is usually unobserved. I provide evidence these assumptions are flawed due to the simultaneity of workplace choice and wage determination, and suggest an alternative identification strategy that exploits previously unused data on average wage at tract of work and local variation in labor demand. I instrument changes in average workplace wage with *local* (tract-level) shift-share labor demand shocks to estimate the labor supply elasticity. Estimates indicate a low value, implying agents are heterogeneous in their preferred locations and relatively unwilling to move in response to changes in local characteristics.[4] A significant advantage to this identification strategy is that it neither requires a major upheaval to city structure (e.g., the division and reunification of Berlin, as in Ahlfeldt et al. 2015) nor requires correct model specification (as in Monte, Redding, and Rossi-Hansberg 2018).

By 2000, LA Metro Rail generates \$109 million-\$146 million in annual surplus. However, these benefits amount to no more than one-half the annualized cost of construction and net operating expenses (depending on the discount rate). Because it may take longer than a decade for commuting behavior to adjust, I draw upon alternative data to test for further changes in commuting after 2000. Tracts first connected before 2000 see an additional increase in commuting of 6%-11% by 2015, and tracts connected after 2000 experience a 12%-13% increase by 2015. Taking these additional gains into account, benefits exceed operational subsidies, but only exceed total costs (including capital expenditures) under a very low discount rate. This analysis suggests that rail transit is unlikely to be cost effective over its first two or three decades as measured by its primary

2. Chen and Whalley (2012) and Kahn (2007) discuss transit-related amenities, while Bowes and Ihlanfeldt (2001) describe disamenities.

3. A standard challenge to hedonic estimates are the presence of price spillovers, which violate the stable unit treatment value assumption (SUTVA). Donaldson and Hornbeck (2016) discuss the importance of modeling general equilibrium when evaluating transportation infrastructure.

4. I also interact local labor demand shocks with the spatial configuration of the city to estimate a tract-scaled housing supply elasticity. Remaining model parameters can taken from the literature or estimated using additional interactions of labor demand shocks with city geography.

output, commuting.[5]

The particular research setting is of great interest: Los Angeles is a large, car-oriented region that built an extensive rail network within a decade. The experience of Los Angeles is more informative for most cities considering rail-based mass transit than evidence from older, denser cities (e.g., Gibbons and Machin 2005). It is an active line of inquiry whether new mass transit infrastructure in less dense cities provides appreciable benefits, particularly given the newer role of cities as centers of consumption in addition to production (Baum-Snow, Kahn, and Voith 2005; Glaeser, Kolko, and Saiz 2001). Interest in understanding the economic consequences of Metro Rail has indeed been high, and there is a budding line of research on the topic.[6]

The paper proceeds to describe the setting in Section 2 and data in Section 3. Section 4 describes identification and estimation of the commuting effect. Section 5 then develops and characterizes the spatial economic geography model. Section 6 discusses the second identification challenge: recovering the elasticities that parameterize the model. I describe estimating the non-commuting effects of transit in Section 7. Section 8 describes welfare estimation, and Section 9 discusses extensions. Section 10 concludes.

2 Setting: Commuting and transit in Los Angeles

High automobile usage has been a long been a feature of commuting in Los Angeles. Angelenos adopted automobiles in large numbers during the rapid growth of the 1920s, leading to early complaints of crowded streets and attempts to relieve traffic delays.[7] Increasing congestion in the 1960s and 1970s led to several failed referendums to expand rail transit.

By 1980, the situation reached a political tipping point. Among the five US Metropolitan Statistical Areas (MSAs) with at least five million residents, Los Angeles residents were the most likely to commute alone in private vehicles and less than half as likely to take transit as the next least transit-intensive MSA. The dominance of the automobile and complex geography prone to bottlenecks meant that Los Angeles consistently ranked as the most congested urban area in the United States. The average trip took one-third longer than the uncongested time, three times the national average delay (Schrank et al. 2015). Los Angeles passed Proposition A in 1980, a sales tax increase partially dedicated to transit. The plan would combine heavy rail (subway) and light rail operations to create an interconnected urban rail transit system. Construction began in 1985.

5. One important caveat is that while I calculate the commuting effects over a 25 year window, I can only examine other channels between 1990 and 2000. Unmeasured benefits include non-commuting trips and environmental factors.

6. Schuetz (2015) shows little change in employment near new Metro Rail stations, and Schuetz, Giuliano, and Shin (2018) ask whether zoning might hinder transit-oriented development near rail stations. Redfearn (2009) studies heterogeneity in the capitalization of the transit amenities in Los Angeles. Anderson (2014) uses transit worker strikes to study congestion spillovers to nearby highways.

7. Chicago had fewer cars entering its urban core in a 24-hour period than Los Angeles did in half a day in the early 1920s, despite Chicago having a population more than twice as large (Kelker, De Leuw & Company 1925).

The first light rail line (the Blue Line) opened in mid-1990 (though construction delays meant it did not reach its urban termini until early 1991). The subway portion experienced several routing changes and first opened in 1993. This line was expanded, with additional stations opening in 1996 and 1999 (the subway is now run as two lines, Red and Purple). Another light rail line initially meant to connect to the international airport (the Green Line) opened in 1995 largely in the median of a new freeway, but was also rerouted.

LA Metro Rail continues to grow. Two lines have already opened and expanded. The Expo Line reached Culver City (another major employment center) in 2012 and now extends to Santa Monica, connecting the system to both the beach and another employment hub. The Gold Line opened in 2003 and now connects downtown LA with areas to the east and southeast. The system currently operates 6 lines, 93 stations, and about 106 miles of rail; current construction will add another line and 17 stations.

3 Data

I develop a panel of tract-level outcomes in 1990 and 2000 that covers Los Angeles County and four adjacent counties (Orange, Riverside, San Bernardino, and Ventura). This five-county area is economically distinct from other conurbations and captures most relevant local interactions. While there is a rich amount of data available, there are some difficulties in obtaining consistent data over the sample period. I briefly discuss my solutions to these issues and data sources; additional details can be found in the Appendix.

Geo-normalization. The standard unit of observation is a census tract or tract pair under 1990 Census geography. Tract definitions change over time, and data products that provide consistent geographies do not include many of the primary variables of interest in this study. I normalize to 1990 geography because it involves the least amount of data manipulation and minimizes rounding issues. I overlay data from more recent geographies and assign to 1990 tracts by coverage, weighting by area when the covarage is split.

Commuting flow data. The primary sources for tract-to-tract commuting flow data are the 1990 and 2000 Census Transportation Planning Packages (CTPP). The CTPP reports aggregate commuting flows between traffic analysis zones, average travel time, some modal information, and various other tabulations. In Los Angeles, traffic analysis zones mostly overlap census tracts; I adjust where necessary. I normalize origin-destination pairs to 1990 geography to create a consistent panel of tract-to-tract commuting flows. Data suppression standards change across CTPP waves, so I apply consistent rounding and suppression rules when combining data across years. In Section 9, I develop a similar dataset covering 2002 and 2015 using LEHD Origin Destination Employment Statistics (LODES), normalized to 2010 geographies. Because of methodological differences in data collection, I do not combine CTPP and LODES.

Place of residence and place of work data. I draw aggregate data on residential census tracts and block groups from the National Historic Geographic Information System (NHGIS). I also use Geolytics' Neighborhood Change Database (NCDB) to validate identifying assumptions. The CTPP contains average *tract of work* wage data unavailable elsewhere, and employment by industry (in 18 aggregate Standard Industrial Classification (SIC) codes). I trim the data to exclude implausible changes between 1990 and 2000 levels (see Appendix for discussion). More recent CTPP products do not include workplace wage, which limits my primary analysis to the 1990-2000 period.

Transit data and treatment; other data sources. I combine geodata on Metro Rail transit stations and lines from the Los Angeles County Metropolitan Transportation Authority (LACMTA) with published information on the timing of station and line openings. To construct labor demand shocks, I draw from IPUMS microdata on all workers outside of California from the 1990 and 2000 Censuses. Panel land use data are from the Southern California Association of Governments (SCAG).

4 Commuting effects of LA Metro Rail

The number of people commuting from residential tract i to workplace tract j at time t, denoted N_{ijt}, depends on residential tract characteristics, θ_{it}, workplace tract characteristics, ω_{jt}, and travel costs τ_{ijt}. Let T denote some function of proximity to transit. Commuting is:

$$N_{ijt} = N_{ijt}\Big(\theta_{it}(T_{it}), \omega_{jt}(T_{jt}), \tau_{ijt}(T_{it}, T_{jt})\Big) \tag{1}$$

The commuting effect of transit captures how connecting i and j changes commuting through travel costs τ_{ijt}. Equation (1) shows that transit can also shift residential or workplace characteristics. Simple regression of commuting flows on transit will not generally differentiate commuting effects from other margins—even if well identified. I discuss identification of the commuting effect of transit, $\frac{\partial N}{\partial \tau}\frac{\partial \tau}{\partial T}$, below, and separately estimate other margins in Section 7.

I rely on both the bilateral and temporal aspects of the flow data. Bilateral data provide a flexible way to control for residential and workplace characteristics and shocks. Temporal variation allows pair-specific fixed effects that control for time-invariant pair characteristics. Let T_{ijt} be a function of T_{it} and T_{jt} that denotes proximity to transit at both origin i and destination j. I estimate:

$$\ln(N_{ijt}) = \omega_{jt} + \theta_{it} + \varsigma_{ij} + \lambda^D T_{ijt} + \varepsilon_{ijt} \tag{2}$$

where ς_{ij} are pair fixed effects and the error captures unobserved, pair-specific shocks to commuting between two locations. Because residential and workplace tract-by-year fixed effects (θ_{it} and ω_{jt}) capture non-commuting effects of transit, λ^D is the average commuting effect of transit.

Equation (2) is a panel gravity equation, where distance is subsumed by (and flexibly controlled for) the time invariant pair fixed effects (e.g., Baier and Bergstrand 2007).[8] I directly model the effect of transit on commuting flows, rather than inferring effects from changes in travel time.

Treatment is defined as proximity of *both a residential and a workplace tract* to LA Metro Rail stations. I use three mutually exclusive, binary definitions of treatment, declining in proximity:

i) *O & D contain station*: Both tracts either contain a transit station or have the centroid within 500 meters of a transit station.

ii) *O & D <250m from station*: Some part of *both* tracts are within 250 meters of a transit station, but at least one tract does not contain, and has a centroid at least 500 meters from, a transit station.

iii) *O & D <500m from station*: Some part of *both* tracts are within 500 meters of a transit station, but at least one tract does not contain, and has a centroid at least 500 meters from, a transit station, and neither tract comes within 250 meters of a transit station.

The median tract is 1.38km^2, so the *O & D <250m from station* bin roughly corresponds to a catchment area of (0km,1.42km] from the station. Only stations open before the end of 1999 are used to define treatment.[9]

In some specifications, I supplement Equation (2) with additional covariates and fixed effects:

$$\ln(N_{ijt}) = \omega_{jt} + \theta_{it} + \varsigma_{ij} + \lambda^D T_{ijt} + \iota_{s_i s_j t} + x'_{ijt}\beta + \varepsilon_{ijt} \tag{3}$$

Subcounty-by-subcounty-by-year fixed effects, $\iota_{s_i s_j t}$, capture regional shifts in commuting patterns. This allows flexible trends in economic integration between pairs of regions (e.g., downtown to west Los Angeles), and limits the variation identifying λ^D pairs of tracts within the same pair of subcounties. Though there are few observable, time-varying, pair specific covariates to include in x_{ijt}, one is potentially important: controls for highway proximity (the Century Freeway opened in the mid-1990s). I do not use variation in travel time: Pair-fixed effects ς_{ij} capture most distance-based variation in times. Further, if travel time responds infrastructure, it becomes an invalid control.

8. One concern is that $N_{ijt} = 0$ for some observations, so $\ln(N_{ijt})$ is undefined. In most specifications, I follow much of the trade literature and exclude pairs with zero flows. There are a few reasons why this is reasonable in my setting. Most pairs connected by transit have non-zero flows, and there is little difference in zero and positive commuting between treated and untreated pairs. I estimate high-dimension fixed effects Poisson PML models for some specifications, and results are qualitatively and quantitatively similar (see Appendix). This is because most pairs that are ever zero (in either 1990 or 2000) are always zero. Always zero pairs do not contribute any variation to models with pair fixed effects; persistent zeros in panel data are less problematic than in the cross-section.

9. The earliest stations opened in July 1990 along the Blue Line, after enumeration of the 1990 Census (in April). The Blue Line did not become fully operational until early 1991 after both endpoint stations opened. Three Red Line stations were completed in each of 1999 and 2000: Stations completed in 1999 are included, those completed in 2000 were finished after Census enumeration and so are excluded.

4.1 Identification

Equation (2) is a bilateral flow analog to difference-in-difference (DD) estimation, supplemented with origin- and destination-by-year fixed effects. Identification requires parallel counterfactual trends: In the absence of treatment, commuting between treated and control tract pairs would have evolved similarly on average, *conditional on separable changes to residential and workplace locations*. This conditioning substantially relaxes standard DD identification. Time-varying origin and destination fixed effects largely control for the non-random siting of transportation infrastructure, as well as other potentially confounding shocks (e.g., to school quality, zoning, etc.)

Identification is instead threatened by the placement of transit routes to connect pairs of locations that would have experienced differential changes in commuting anyway. I limit selection in route placement by drawing on historical data giving the locations of a proposed subway network and former streetcar lines and shocks to route placement. These locations select pairs of tracts that could have plausibly received transit by 2000 and that share common historical land use and transportation characteristics that may still influence urban outcomes today. Thus, commuting between selected locations likely evolves in a similar counterfactual manner.[10]

I draw from Kelker, De Leuw & Company (1925), which details a feasible rail transit network designed to accommodate Los Angeles' booming population in the 1920s. The plan was defeated largely because of skepticism over private rail management and local opposition to elevated portions of the line.[11] This document also shows Pacific Electric Railroad (PER) lines installed in 1925. The PER, colloquially called Red Cars, was an at-grade railway system that served Los Angeles through 1961. I define two samples as the union of pairs of tracts near LA Metro Rail by 2000 and pairs that lie within 1km of: (i) the Kelker, De Leuw & Company subway proposal, "1925 Plan Sample"; or (ii) PER lines, "PER Sample". The 1925 Plan Sample itself has two variants: an immediate plan meant to be built right away, and the full plan meant to accommodate buildout. Maps from Kelker, De Leuw & Company (1925) are shown in Figure 1; the subway lines, stations, and treated and control tracts are mapped in Figure 2 (subcounties are outlined in black).

The validity of these groups as controls is supported by several lines of reasoning and evidence. First, many control pairs contain one 'end' (either the origin or destination) that is treated, though the other end is not. Such control pairs compare changes in the commuting behavior of residents of i who work in j (which receives a transit linkage) and those who work in j' (which does not). Similarly, workers in j who reside in i are compared with those who reside in i'. These comparisons control for many potential unobserved motives for changing commuting behavior.

10. Following Baum-Snow (2007), researchers increasingly use historical plans for exogenous variation in infrastructure placement. I use historical data to select controls rather than instrument for route location, primarily because Brooks and Lutz (2016) find path dependent effects of historical transit infrastructure. Their finding invalidates historical networks as instruments within cities, but does suggest their use to define matched control groups.

11. The transit system was to be run by Southern Pacific Railroad, which had a significant (and perhaps overlarge) influence on regional politics (Fogelson 1967). Many of these alignments are part of modern LA Metro Rail.

Second, there was significant variation in timing and prioritization of route construction, due largely to reasons orthogonal to transit demand. Most notably, a geologic shock delayed the westward expansion of the Red/Purple Line by roughly 30 years. The original routing for this line was to run along Wilshire Boulevard, one of the densest corridors in Los Angeles. Methane seepage into a nearby Ross Dress for Less store exploded on March 24, 1985, leading to federal legislation restricting tunneling along Wilshire. This corridor appears in almost every transit plan from the 1920s until today; both 1925 Plan and PER Samples select it as a control. The Green Line route was chosen to minimize construction costs by lying partially within an under-construction highway. The westward end was first meant to connect to Los Angeles International Airport (LAX), but concerns about electromagnetic interference raised by the Federal Aviation Administration disrupted this alignment, leading to a more southerly route. A connection to LAX is now under construction, and many other tracts selected as controls have become treated as the network expands.

Route determination was also impacted by other factors besides transportation patterns.[12] Routes were designed to satisfy political pressures, ensure political support for allocating revenue to rail projects, and spur political favor for one of the two agencies overseeing rail in Los Angeles. To illustrate, politicians demanded that heavy rail serve the San Fernando Valley, despite the cost and difficulty of doing so. It was also deemed necessary to connect Long Beach to ensure its portion of state gas tax revenue could be diverted to rail funding. At one point, a serpentine route under consideration was dubbed the 'wounded knee' because its illogical shape touched many local political jurisdictions. Furthermore, the two agencies overseeing rail had different motivations and different supporters. The Los Angeles County Transportation Commission promoted light rail and a broad service area, while the Southern California Rapid Transit District promoted heavy rail downtown and along dense corridors. The two agencies half-cooperated, half-competed for limited funds, each seeking the lead role in a viable system. In sum, "politics, outside circumstances, and the geography of power … played an outsized role in influencing where the new rail lines would go" (Elkind 2014, p. 50).

Finally, I provide econometric evidence to support the choice of control pairs. Commuting data before 1990 are unavailable, so I cannot provide direct evidence of parallel pre-trends in *tract-pair* flows. However, I compare pre-trends of *tract*-level housing and labor market characteristics using NCDB data from 1970 to 1990.[13] The primary economic variables used in the model show little evidence of differential pre-trends in between treatment and control groups: Employment, households, income, and housing values are generally on similar trends (Appendix Table F1). Transportation behaviors between treated and control tracts are also generally on parallel pre-trends, though tracts that become treated show slightly more transit use (Appendix Table F2). Generally,

12. These arguments draw in part on the history LA Metro Rail in Elkind (2014), who notes that "plotting a subway through the politically decentralized landscape of Los Angeles meant ceding control to numerous fiefdoms of federal, state, and local politicians" (p. 70).

13. NCDB is, by default, normalized to 2010 geographies. I use the same treatment rules, but this results in higher observations counts due to denser tracts in 2010 than 1990.

pre-trends in tracts appear parallel. Furthermore, any tract-level differences are absorbed by the tract-by-year fixed effects; unobserved pre-trends in bilateral flows are likely parallel.

4.2 Commuting flow estimates

LA Metro Rail led to an average increase of 10%-16% in commuting between tracts nearest transit stations by 2000 (Table 2). Results are significant across control group specifications, and robust to the inclusion of subcounty pair-by-year fixed effects and controls for highway proximity. Only tracts containing or within 250m of a station show a significant effect. Results are largest using the 1925 Subway Plan control groups, varying between 11% and 16%. Columns 1-5 use a 'loose' network, in which tract pairs with one treated and one control tract are retained. Column 6 uses a 'tight' network that excludes such pairs. This distinction makes little difference in the 1925 Plan samples, but 'tight' estimates shrink in the PER Sample. Standard errors are clustered along three dimensions to be robust to correlation within tract pairs, residential tracts, and workplace tracts.

Columns 1 to 3 add in the different measures of treatment intensity, and suggest the three-bin defintion captures the effect reasonably.[14] The results in Columns 4 and 5 further include subcounty pair-by-year fixed effects to account for secular interregional changes in commuting patterns. These fixed effects increase the size of most estimated effects. Finally, Column 5 adds controls to account for the opening of the Century Freeway in southern Los Angeles County. Inclusion decreases point estimates by a small amount, but estimates of the effect of transit on commuting flows are still substantial and significant.

The rich commuting flow data allow further identification checks and explorations of effect heterogeneity. Appendix Table F5 compares the effect of being connected for tract pairs *on the same line* with tract pairs connected but not on the same line. Thus both routes become treated, but only for tract pairs on the same line is there the most substantial reduction in travel costs. The difference in these effects for pairs that both contain stations (10%-13%) is similar in size to those in Table 2. Interacting proximity measures for origin and destination tracts indicate a greater effect of proximity at the destination than the origin. This suggests commuters are more comfortable traveling a larger distance to a rail station from home than from a rail station to work. I do not find evidence of heterogeneous treatment effects by origin-destination distance.

4.3 Commuting time estimates

A motivation for LA Metro Rail was to relieve automobile congestion. I use changes in reported CTPP travel times to test whether transit decreases congestion:

$$\tau_{ijt} = \omega_{jt} + \theta_{it} + \varsigma_{ij} + \lambda^{\tau,2\text{km}} 1_{ij \text{ within 2 km},t} + \lambda^{\tau,4\text{km}} 1_{ij \text{ 2 to 4 km},t} + \varepsilon_{ijt} \tag{4}$$

14. After much experimentation, this approach seems to best balance parsimony and completeness.

where τ_{ijt} is the average reported travel time from i to j in year t. I use two mutually exclusive indicators for whether the pair lies within a 2km or 4km corridor of the subway system. Ideally, I would observe which pairs require driving routes that are potentially affected by transit. Wide corridors in Equation 4 capture routes most likely impacted, but may miss longer routes that bisect transit corridors.

Table 3 shows results for three measures of travel time: average travel time across all modes, its log, and average travel time for private cars. For the first two measures, I include a control for immediate transit station proximity so as to capture confounding travel time changes for rail commuters. Across all three measures, simpler models yield significant decreases of 1.3-1.4 minutes, or about 3.2%. Results remain negative but are insignificant with subcounty pair-by-year fixed effects. While this suggests that there is no congestion effect, the fixed effects may absorb potentially useful variation. I find no evidence of increased commuting within the 2km bandwidth.

These results are informative about long run attenuation of the results in Anderson (2014), who finds that a temporary labor strike disrupting LA Metro Rail service in 2003 increased nearby automobile congestion. Travel demand is particularly responsive to short run changes in congestion, so it is unclear how to map temporary to long run responses. For example, Duranton and Turner (2011) find no aggregate evidence that transit decreases automobile travel.[15] When significant, the results in Table 3 are roughly one-quarter to one-third those in Anderson (2014). This suggests substantial attenuation of congestion benefits.

5 A model of urban location choice

To translate the effects of transportation infrastructure to welfare, I describe a quantitative urban model of residential and workplace choice with commuting. The model links local, observable equilibrium outcomes to local, unobservable economic fundamentals (e.g., productivity, amenities, housing efficiency). The model includes a collection of N locations in a city, operationalized as census tracts, that each contain a labor market and a housing market. Agents choose the location pair that maximizes utility.

The model is similar to that of Ahlfeldt et al. (2015), with five differences. The first two generalize Ahlfeldt et al. (2015): (i) origin-destination pairs can differ in mean utility and (ii) a local housing efficiency parameter captures differences in local regulations and per unit housing costs. Two further differences are simplifications that match the empirical setting and have little quantitative impact: (iii) land use between housing and production is exogenously determined and (iv) agglomeration and consumption externalities are excluded from the primary model.[16] Differ-

15. At play is the 'Fundamental Law of Congestion' (Downs 1962). If this holds, then any improvements to congestion, or air pollution, are transitory, and the primary role of transit is to enable a larger population. Empirical evidence generally supports this 'law' (e.g., Duranton and Turner 2011; Hsu and Zhang 2014).

16. I later document that there is little scope for land use adjustment. Prior versions of this paper included spillovers,

ences (i) to (iii) lead to difference (v): the model can be transparently rewritten as a system of three equations log-linear in available data.

Joint market household decision: Labor supply and housing demand

Atomistic households make location and consumption decisions. For the location decision, households choose a tract of work and a tract of residence. Conditional on choosing to live in location i, households face per unit housing costs Q_i and receive amenity $\tilde{B}_i$. Conditional on choosing place of work j, households inelastically provide one unit of labor in exchange for wage W_j. Given the joint location choice and prices, households make decisions over consumption of housing and a composite good. Specifically, household o chooses location pair ij, consumption $\mathcal{C}$, and housing H to maximize the following Cobb-Douglas utility function:

$$\max_{\mathcal{C},H,\{ij\}} U_{ijo} = \max_{\mathcal{C},H,\{ij\}} \frac{\nu_{ijo}\tilde{B}_i}{\delta_{ij}} \left(\frac{\mathcal{C}}{\zeta}\right)^{\zeta} \left(\frac{H}{1-\zeta}\right)^{1-\zeta} \quad \text{s.t.} \quad \mathcal{C} + Q_i H = W_j$$

where ν_{ijo} is household o's idiosyncratic preference for location pair ij. The cost of commuting between i and j is captured by $\delta_{ij} = e^{\kappa\tau_{ij}} \geq 1$, where τ_{ij} is a measure of travel cost. The share of household expenditures on housing is $1-\zeta$. Indirect utility conditional on location pair ij is:

$$v_{o|ij} = \frac{\nu_{ijo}\tilde{B}_i W_j Q_i^{\zeta-1}}{\delta_{ij}}$$

Given this specification, optimal housing consumption for household o conditional on location pair ij is given by $H_{ijo} = (1-\zeta)W_j/Q_i$.[17]

To map indirect utility to choice probabilities, assume ν_{ijo} is distributed Fréchet with scale parameter $\tilde{\Lambda}_{ij} = T_i E_j D_{ij}$ and shape parameter $\epsilon > 0$. The cdf of ν is thus:

$$F_{ij}(\nu) = e^{T_i E_j D_{ij} \nu^{-\epsilon}}$$

The scale parameter captures mean idiosyncratic preference for location pair ij: T_i captures the mean utility of residing in i, E_j the mean non-wage utility of working in j, and D_{ij} an unobserved pair-specific shift in the utility of a particular commute. The shape parameter governs the degree of homogeneity in preferences: For high ϵ, agents view location pairs homogeneously, while for low ϵ, their valuations are heterogeneous. With this distributional assumption, utility maximization yields a simple proportional formula for commuting flows. The share of the population that

however, they were quantitatively insignificant (they are essentially time invariant and thus subsumed by tract fixed effects). I explore relaxing these conditions in Section 6 and the Appendix.

17. Any indirect utility function with a multiplicatively separable idiosyncratic component could be employed. For example, the sorting literature uses a nested CES parameterization (Epple and Sieg 1999). Davis and Ortalo-Magné (2011) show that expenditure shares on housing are relatively constant through time in cities in the United States, supporting the Cobb-Douglas assumption. I retain this assumption to maintain comparability with the existing literature.

chooses residential location i and place of work j is:

$$\pi_{ij} = \frac{\tilde{\Lambda}_{ij}\left(\delta_{ij}Q_i^{1-\zeta}\right)^{-\epsilon}(\tilde{B}_i W_j)^{\epsilon}}{\sum_r \sum_s \tilde{\Lambda}_{rs}\left(\delta_{rs}Q_r^{1-\zeta}\right)^{-\epsilon}(\tilde{B}_r W_s)^{\epsilon}} \tag{5}$$

To relate commuting shares to observable commuting flows, multiply π_{ij} by the population of the market as a whole ($\bar{N}$), so that $N_{ij} = \pi_{ij}\bar{N}$.

The city can be viewed either as existing in autarky or being nested in a large, open economy. This assumption makes little difference outside of welfare calculations (due to homothetic preferences). In an open economy, no spatial arbitrage requires that the average welfare from moving to the city equal the reservation utility of living anywhere else. The expected value of moving to the city is:

$$\mathbb{E}[U_{ijo}] = \Gamma\left(\frac{\epsilon - 1}{\epsilon}\right) \cdot \left[\sum_r \sum_s \tilde{\Lambda}_{rs}\left(\delta_{rs}Q_r^{1-\zeta}\right)^{-\epsilon}(\tilde{B}_r W_s)^{\epsilon}\right]^{1/\epsilon} \tag{6}$$

where $\Gamma(\cdot)$ is the gamma function and the aggregate population $\bar{N}$ is implicitly defined. Free mobility thus requires $\mathbb{E}[U_{ijo}] = \bar{U}$, and aggregate population changes to maintain $\bar{U}$.

Production: Labor demand

A continuum of measure zero firms produces a globally tradable commodity in each location j under perfect competition. Firms select competitively available labor N^Y and land L^Y inputs to maximize profits under constant returns to scale.[18] Production is multiplicatively separable in local productivity A_j and a technology that is identical across j:

$$Y = A_j F(N_j^Y, L_j^Y) \tag{7}$$

Because of the atomistic size of firms, land use decisions are made in accordance with profit maximization despite the locally fixed available quantity of land.[19] Perfect competition in labor markets implies that firms pay workers the marginal product of labor: $W_j = A_j F_N(N_j^Y, L_j^Y)$. I assume Cobb-Douglas production technology: $F(N^Y, L^Y) = (N^Y)^{\alpha}(L^Y)^{1-\alpha}$. Inverse labor demand is given by:

$$W_j = \alpha A_j \left(\frac{L_j^Y}{N_j^Y}\right)^{1-\alpha} \tag{8}$$

18. The primary focus of this study is the flow of people rather than the flow of goods, so I assume that goods are uniformly available and globally traded.

19. Individual firms make unconstrained input decisions, but aggregate land use is predetermined, as is standard in many urban models, e.g., Glaeser et al. (2008).

Housing supply

Housing is produced by measure zero builders using land for housing L^H and material inputs M. A local, multiplicatively separable housing productivity term $\tilde{C}_i$ captures local cost drivers such as geography (e.g., terrain) and regulation. Materials are readily available in all locations at the same cost, but aggregate local land supply for housing is predetermined.[20] Convexity in land pricing serves as a congestive force, driving up prices in desirable locations until agents look elsewhere. I specify Cobb-Douglas housing production: $H = (L^H)^{\phi} M^{1-\phi} \tilde{C}_i$.[21] Developers sell housing in location i in a competitive market at unit price Q_i to maximize profit: $Q_i H - P_i^L L^H - P^M M$. The price of construction materials P^M is exogenous and common to all locations.

Because detailed data on housing production are not available, I utilize the zero profit condition to develop an empirical formula for housing costs. The first order condition of developer profit with respect to construction materials gives:

$$Q_i = \frac{P^M}{(1-\phi)\tilde{C}_i} \left(\frac{M}{L^H} \right)^{\phi} \tag{9}$$

Substituting this into the developer's profit function and enforcing the zero profit condition implied by perfect competition gives construction material demand: $M^* = \frac{1-\phi}{\phi} \frac{L^H P_j^L}{P^M}$. Enforcing zero profits gives $Q_i = (P_i^L L^H + P^M M)/((L_i^H)^{\phi} M^{1-\phi} \tilde{C}_i)$. Substituting in M^* gives the cost function: $Q_i = C_i (P_i^L)^{\phi}$, where $C_i = (P^M)^{1-\phi}/(1-\phi)^{1-\phi} \phi^{\phi} \tilde{C}_i$ captures the inverse efficiency in housing production.

The price of land, P_i^L, responds to changes in demand and land availability: I parameterize it as a function of local housing density $P_i^L = (H_i/L_i^H)^{\tilde{\psi}}$, where the parameter $\tilde{\psi} > 0$ captures local price elasticity of land with respect to density.[22] This parameter provides a congestive force to the model. Combining the expression for land price with Equation (9) and compressing notation relates housing supply, price, and land availability:

$$Q_i = C_i \left(\frac{H_i}{L_i^H} \right)^{\psi} \tag{10}$$

where $\psi = \tilde{\psi}\phi$. As housing productivity $\tilde{C}_i$ increases, C_i falls, so increases in housing productivity (decreases in C_i) increase the quantity of housing supplied at any price.

20. This simplifies the model while maintaining fidelity to the setting. Strong zoning and the medium time frame of this study may not match the temporal patterns required for land use change; many studies of land use or housing supply examine only long-run changes (e.g., Saiz (2010) uses a 30 year window). Including land use measures does not greatly change identification or results. There is little evidence of differential land use near transit.

21. Ahlfeldt et al. (2015), Combes, Duranton, and Gobillon (2012), and Epple, Gordon, and Sieg (2010) show that Cobb-Douglas works well for floor space production with land and material inputs in several settings.

22. I discuss an alternate way to close the model in the Appendix. Because I have data on land use, rather than floor space or use, I frame the model and analysis in terms of land.

Equilibrium characterization

In equilibrium, labor and housing markets clear in all locations. Labor market clearing requires that local labor demand equal supply:

$$N_i^Y = \sum_r \bar{N}\pi_{ri} \tag{11}$$

Commuting shares (1) determine employment in any location. With frictional commuting, workers benefit from residing near work locations. Given the assumption on preferences, housing demand is a constant fraction of the ratio of wage to housing price. Aggregate housing demand in i is the sum of wage-rent ratios weighted by commuting flows—this takes into account heterogeneity in income stemming from variation in place of work. Housing market clearing requires that the local housing supply equal demand:

$$H_i = (1-\zeta)\sum_s \bar{N}\pi_{is}\frac{W_s}{Q_i} \tag{12}$$

Given model parameters $\{\alpha, \epsilon, \zeta, \psi, \kappa\}$, reservation utility $\bar{U}$, vectors of land availability by use $\{\mathbf{L^Y}, \mathbf{L^H}\}$, vectors of residential fundamentals $\{\tilde{\mathbf{B}}, \mathbf{C}, \mathbf{T}\}$, vectors of place of work fundamentals $\{\mathbf{A}, \mathbf{E}\}$, and matrices of residential-place of work pair fundamentals $\{\mathbf{D}, \boldsymbol{\tau}\}$, an equilibrium is referenced by price vectors $\{\mathbf{W}, \mathbf{Q}\}$, commuting vector $\boldsymbol{\pi}$, and scalar population measure $\bar{N}$.

Proposition 1. *Consider the equilibrium defined by equations (5), (8), (10), (11), and (12):*

i) At least one equilibrium exists across residential locations with strictly positive quantities of residential land and work locations with strictly positive quantities of land used in production.

ii) There is at most one equilibrium if

$$\frac{2\epsilon(\epsilon+1)(1-\alpha)(1-\zeta)}{1+\epsilon(1-\alpha)} - 1 \leq \frac{1}{\psi} \tag{13}$$

Proof. See Appendix.[23] □

Existence makes use of the assumption that land use is predetermined and requires that positive residential land translates to a positive measure of residents and that positive land in production translates to a positive measure of workers. However, existence does not require positive commuting flows between all locations. The presence of zero commuting flows is a common characteristic of commuting data. The uniqueness condition requires that the elasticity of housing

23. Existence makes use of the Brouwer fixed point theorem; uniqueness relies on Theorem 1 in Allen, Arkolakis, and Li (2014) and the Perron-Frobenius theorem.

supply ($1/\psi$) be larger than a function of preference homogeneity and other parameters. The left-hand term is increasing in ϵ: The more homogeneous preferences are, the more elastic housing supply must be to ensure a single equilibrium.

Inversion

The model may have multiple equilibria (though this is unlikely given the low value of ϵ in Section 6). Regardless, for a given set of parameters, there is a unique mapping from the observed data to local fundamentals. Model parameters are estimated using these fundamentals and the observed values of the endogenous variables in combination with instruments to define moment conditions. $\tilde{B}_i$ and T_i enter isomorphically; let $B_i = T_i \tilde{B}_i^{\epsilon}$ and $\Lambda_{ij} = B_i E_j D_{ij}$.[24] Local fundamentals $\mathbf{A}$, $\mathbf{C}$, and $\mathbf{\Lambda}$ can be expressed as unique functions of data and parameters:

Proposition 2. *Given parameters $\{\alpha, \epsilon, \zeta, \psi, \kappa\}$, observed data $\{\mathbf{W}, \mathbf{Q}, \boldsymbol{\pi}, \bar{N}\}$, and commuting times $\boldsymbol{\tau}$, then there exists a unique set of fundamentals $\{\mathbf{A}, \mathbf{C}, \mathbf{\Lambda}\}$ that are consistent with the data being an equilibrium of the model.*

Proof. See Appendix. □

6 Identification and estimation

Local labor and housing market elasticities provide the mapping between local fundamentals (and interventions that shift them) and observed prices and quantities. Consistent estimates of the elasticities are required to use observable data to learn about changes to local fundamentals and to simulate counterfactual scenarios. I develop an identification strategy that uses panel variation in wages at place of work, housing prices, and commuting flows, permitting the incorporation of tract and tract-pair fixed effects to flexibly control for unobserved, time-invariant characteristics that confound identification (for example, location near a port, or on a hillside with a nice view, or in town with very stringent land use regulations).[25]

All components of the model are expressed in the commuting flow (1), wage setting (8), and housing price (10) equations. Denote by lowercase letters log values. Adding time subscripts and including tract and tract-pair fixed effects (so that, e.g., $\ln(A_{jt}) = \bar{a}_j + a_{jt}$), I show in the Appendix

24. This mapping diverges from Ahlfeldt et al. (2015), where local fundamentals consist of a composite workplace term that combines $\mathbf{A}$ and $\mathbf{E}$, a residential term that combines $\mathbf{B}$ and $\mathbf{T}$, and omits location or pair specific variation in housing supply $\mathbf{C}$ or commute utility $\mathbf{D}$. The components of $\mathbf{\Lambda}$ are not uniquely identified from the data; I use statistical arguments to separate $\mathbf{B}$, $\mathbf{E}$, and $\mathbf{D}$.

25. Persistent, difficult-to-measure characteristics play an anchoring role in cities (Lee and Lin 2018). Land use regulation locks in such differences in amenities in Southern California (Kahn, Vaughn, and Zasloff 2010; Severen and Plantinga 2018).

that these equations deliver a tractable system log-linear in data and fundamentals:

$$\begin{aligned}
\text{Labor demand:} \quad & w_{jt} = g_{0t} + \tilde{\alpha} n^Y_{jt} + \bar{a}_j + a_{jt} && (14)\\
\text{Commuting:} \quad & n_{ijt} = g_{1t} + \underbrace{\epsilon w_{jt} + \bar{e}_j + e_{jt}}_{\substack{=\,\omega_{jt},\\ \text{Labor supply}}} + \underbrace{\tilde{\zeta} q_{it} + \bar{b}_i + b_{it}}_{\substack{=\,\theta_{it},\\ \text{Housing demand}}} - \epsilon\kappa\tau_{ijt} + \bar{d}_{ij} + d_{ijt} && (15)\\
\text{Housing supply:} \quad & q_{it} = g_{2t} + \psi h_{it} + \bar{c}_i + c_{it} && (16)
\end{aligned}$$

where n^Y_{jt} is log employment density, h_{it} is log housing density, the g are constants, and $\tilde{\alpha} = \alpha - 1$ and $\tilde{\zeta} = -\epsilon(1-\zeta)$. This system is a within city analog of the Roback (1982) and Rosen (1979) framework with commuting. Local fundamentals are potentially functions of covariates ($\bar{a}_j + a_{jt} = a(x_{it})$ and so on), such as transit proximity. Equation (15) provides structural interpretation of empirical Equation (2).

6.1 A general approach to identifying local elasticities

I develop a local implementation of a shift-share (e.g., Bartik 1991) instrument to overcome simultaneity in Equations (14) to (16). I leverage plausibly exogenous panel variation in tract-level labor demand, interacting local labor demand shocks with the distance between tracts to create exogenous variation in local economic conditions. I focus on identification of ϵ (the elasticity of labor supply) and ψ (the inverse elasticity of housing supply) in the paper, as these two embed information about the local economic environment and cannot be estimated from microdata.[26] I also discuss the robustness of this strategy to alternative assumptions.

Identification requires a demand or supply shock that shifts one of Equations (14)-(16) but is excludable from the others. I construct tract-level labor demand shocks from changes in national wage and employment levels and ex ante local employment shares by industry. After controlling for year and census tract fixed effects, the remaining variation consists of changes in wages and employment determined from ex ante, local industrial composition. These shocks are relevant if they are correlated with changes in local productivity (Δa_{jt}) and excludable if they are uncorrelated with changes in the other local fundamentals. Under these assumptions, the labor demand shock traces out the labor supply curve. Housing demand in nearby locations shifts in response. Because this downstream housing demand response will be stronger nearer the workplace origination of the shock, I take a linear combination of labor demand shocks with weights determined by a spatial decay function and commuting to map the labor demand shocks to a residential tract. This derived housing demand instrument traces out the housing supply curve.[27]

26. I develop moment conditions can identify all four housing and labor supply and demand elasticities in the Appendix. However, these assumptions required to identify the demand elasticities are stronger. While I recover reasonable estimates for the demand elasticities, values can be easily taken from the literature or estimated from microdata.

27. The Appendix provides details on the identifying assumptions for the demand elasticities. To briefly summarize, identifying housing demand requires an instrument that shifts housing supply. For agents who work in j and live in i,

Let $R_t^{q,Nat}$ be average national wage or total national employment in industry q in year t, $N_{j,0}^q$ be the number of workers in each two-digit SIC industry q in the initial year (1990) in tract j, and $N_{i,0} = \sum_q N_{j,0}^q$ the ex-ante total employment in tract i. The labor demand shock is formed by interacting changes in wages or employment with ex ante local employment shares and summing across industries:

$$\Delta z_{jt} = \sum_q \frac{R_t^{q,Nat} - R_0^{q,Nat}}{R_0^{q,Nat}} \cdot \frac{N_{j,0}^q}{N_{j,0}}$$

This demand shock embeds information on ex ante industry shares. When used as an instrument, an implicit assumption is that changes in non-productivity latent variables (e.g., amenities) are uncorrelated with prior industry structure.[28] To ensure that local factors do not drive national changes, I exclude all workers in California.

6.2 Identifying the labor supply elasticity (Fréchet shape parameter)

The Fréchet shape parameter ϵ governs the homogeneity of location preference, but is also an extensive margin labor supply elasticity that conditions on commuting and residential geography. A straightforward approach to identify the slope of labor supply is to instrument the time-varying wage in Equation (15) with the demand shock. Place of residence-by-year fixed effects (θ_{it}) control for changes in residential amenities that may be correlated with labor demand shocks. The corresponding moment condition is:

$$\mathbb{E}[\Delta z_{jt} \times (\Delta e_{jt} + \Delta d_{ijt})] = 0, \ \forall i, j \tag{M-1}$$

This condition can be weakened by first estimating place of work-by-year fixed effects to use as the dependent variable measuring labor supply conditional on commuting. From Equation (15), $\omega_{jt} = \epsilon w_{jt} + \bar{e}_j + e_{jt}$. Instrumenting wage with the labor demand shock in this setting identifies the labor supply elasticity if

$$\mathbb{E}[\Delta z_{jt} \times \Delta e_{jt}] = 0, \ \forall j. \tag{M-1a}$$

Moments conditions M-1 and M-1a require that the constructed local labor demand shock in a tract j be uncorrelated with unobservable changes in factors that shift labor supply to that same tract. These unobservable factors are changes in workplace amenities (or any workplace-specific

a labor demand shock to agents who also live in i but work elsewhere (in j') shifts effective housing supply in i. That is, a labor demand shock for workers $n_{ij'}$ with $j' \neq j$ translates into a housing supply shock to workers n_{ij} (as long as housing supply is not perfectly elastic). I again use spatial decay weights and commuting to determine an appropriate relationship between labor demand shocks and the housing market. Finally, labor demand shocks in one location alter wages and induce workers to shift employment location.

28. A recent literature considers the consequences of this in detail (e.g., Adão, Kolesár, and Morales 2019; Goldsmith-Pinkham, Sorkin, and Swift 2018).

labor supply shifter), e, and travel costs, d.

Results and discussion

I proceed to estimate ϵ under moment condition M-1a, first recovering workplace-by-year fixed effects ω_{it} from the Equation 15, then instrumenting the change in wage $\omega_{jt} = \epsilon w_{jt} + \bar{e}_j + e_{jt}$ with the labor demand shock Δz_{jt}. I use the wage variant of the labor demand shock, as wage is the endogenous variable.

Table 4 uses three different methods of estimating ω_{it}. Column 1 estimates ω_{it} jointly using both years of data in a log-linear panel, and so this measure of ω_{it} flexibly controls for geography through bilateral time-invariant fixed effects d_{ij}. Column 2 uses a separate PPML estimator for each year to estimate ω_{it}, conditioning on bilateral travel costs. Column 3 jointly using both years of data like Column 1, but uses panel PPML estimator with bilateral time-invariant fixed effects flexibly control for geography. The first stage is sufficiently strong across all specifications. The value of ϵ is 0.50 under the linear specification, and between 1.83 and 1.85 in using the nonlinear estimator.[29] I take $\epsilon = 1.83$ from Column 3 as the preferred estimate. The low value of ϵ implies workers are quite heterogeneous in their location preferences.

This estimate of ϵ is roughly in line with extensive-margin labor supply elasticities found in studies of labor markets, and substantially smaller than most values in the urban economic geography literature. For example, Falch (2010) estimates labor supply elasticities between 1.0 and 1.9, Suárez Serrato and Zidar (2016) find values between 0.75 and 4.2, and Albouy and Stuart (2014) recover 1.98. In fact, conditions M-1 and M-1a are local, rather than citywide, statements, and are substantially weaker than standard applications of shift-share instruments. When an aggregate (citywide) labor demand shock is used to trace out aggregate labor supply, identification requires the shock be orthogonal to any non-wage determinants of labor supply (residential amenities, commuting costs, workplace amenities). In my notation:

$$\mathbb{E}[\Delta \bar{z}_t \cdot (\Delta \bar{b}_t + \Delta \bar{d}_t + \Delta \bar{e}_t)] = 0$$

where $\bar{\cdot}$ averages over locations within a city. This suggests potential pitfalls in standard applications of shift share instruments. First, changes in residential amenities, commuting costs, and workplace amenities must be uncorrelated with labor demand shocks *regardless of where in the city they occur*. Second, changes in amenities cannot be correlated with the labor demand shock. This concern leads Diamond (2016) to both measure amenity levels and interact city-level labor demand shocks with housing regulations and geography. In my setting, even unobserved residential amenities are not confounding. Finally, changes in unobserved commuting costs cannot

29. Unlike the commuting analysis, accounting for zeros makes a large difference in estimates. This is because any individual work tract has many zeros, and so incorporating this information into the fixed effect is important. See Larch et al. (2019) for a discussion of the estimator.

be correlated with the labor demand shock locally or elsewhere within the city. This would be violated if, e.g., growth in trucking created significant congestion. In contrast, moment conditions M-1 and M-1a clarify the spatial requirements for identification and are robust to both local and non-local correlation between improvements in residential amenities or commuting costs.

The urban economic geography literature typically identifies ϵ from a combination of modeling assumptions and cross-sectional variation in travel time (or distance), often resulting in higher estimates of ϵ. For example, Ahlfeldt et al. (2015) condition on the marginal disutility of travel time estimated from auxiliary models, then require two cross-sectional assumptions: wage must be orthogonal to workplace wage ($\mathbb{E}[w_j e_j] = 0$) and there can be no variation in (non-pecuniary) workplace utility ($\mathbb{E}[e_j^2] = 0$).[30] Such assumptions not supported by the data. Panel A of Figure 3 plots ω_j (which is equal to $\epsilon w_j + e_j$) against w_j using data from Los Angeles in 1990. If $\mathbb{E}[w_j e_j] = 0$, then wages and labor supply are not simultaneously determined, and the slope in Panel A of Figure 3 (0.17) is equal to ϵ. If $\mathbb{E}[e_j^2] = 0$, this figure would just reveal a one-to-one mapping between ω_j and w_j. Instead, ω_j is more closely related to workplace employment levels (Panel B of Figure 3), highlighting the severity of simultaneity when wage is unobserved.[31] This value of ϵ I estimate has important implications for studies of urban structure: Preference heterogeneity limits the responsiveness of agents to changing local conditions.

6.3 Identifying the (inverse) housing supply elasticity

A labor demand shock in one location shifts demand for housing in locations where workers might live, and thus can be used to instrument changes in housing quantity to identify the slope of the housing supply curve. This requires mapping the labor demand shock to residential locations. I describe a housing shock to residential location i of the form $\Delta z_{it}^{HD} = \mathbf{z}_t \cdot \boldsymbol{\vartheta}_i$, and treat weights ϑ as a parametric decay functions of travel time between locations that ever have positive commuting:

$$\Delta z_{it}^{HD}(\rho) = \sum_s \frac{e^{-\rho \delta_{is}} 1_{\check{n}_{is}>0} \Delta z_{st}}{\sum_s e^{-\rho \delta_{is}} 1_{\check{n}_{is}>0}}$$

30. Ahlfeldt et al. (2015) generally allow for more flexible e_j, but make the stricter assumptions to identify the Fréchet shape parameter. To see this, Equation 35 in Ahlfeldt et al. (2015) requires $\mathbb{E}[(1/\epsilon)^2 \omega_j^2 - \sigma_w^2] = 0$, where $\omega_j = \epsilon w_j + e_j$ and σ_w^2 is observed wage dispersion. Rearranging this condition yields identification of ϵ if $\epsilon^2 \mathbb{E}[w_j^2] + 2\epsilon \mathbb{E}[w_j e_j] + \mathbb{E}[e_j^2] = \epsilon^2 \mathbb{E}[w_j^2]$.

31. Other papers take varied, ad hoc approaches to deal with unobservable workplace wage. Allen, Arkolakis, and Li (2015) assume that marginal disutility of travel time is equal to the marginal utility of working. Monte, Redding, and Rossi-Hansberg (2018) assume an elasticity of substitution σ, specify a trade-in-goods model to recover productivity from cross-sectional trade flows, then assume recovered productivity is orthogonal to workplace and origin-destination specific amenities $\mathbb{E}[a_j(\sigma) \times (e_j + d_{ij})] = 0, \forall i, j$. This requires correct model specification to recover a_j. Tsivanidis (2018) proceeds more carefully, first estimating the marginal disutility of travel time estimated from a model of mode choice, and running panel models similar to Equation (15) and exploiting exogenous variation in transit connections. Exceptionally, Kreindler and Miyauchi (2017) observe workplace wage and find a modest cross-sectional relationship between ω and w using cell phone data and travel surveys from Dhaka, Bangladesh and Colombo, Sri Lanka.

where δ_{js} is the travel time between j and s, ρ is the spatial decay parameter, and $\check{n}_{is}$ denotes the maximum flow value from i to s in any year. These weights dictate that labor demand shocks nearer a residential location with a positive commuting connection are more important than labor demand shocks farther from i or in places with no commuting connection to i. The resulting inverse-distance weighted labor demand shock can be used to instrument housing density and identify ψ, the inverse price elasticity of housing supply, under the following condition:

$$\mathbb{E}[\Delta z_{it}^{HD}(\rho) \times \Delta c_{it}] = 0, \ \forall i \tag{M-2}$$

Although both elements of M-2 relate to tract i, the housing demand shock draws on labor demand shocks from any j (including i).

Moment condition M-2 requires labor demand shocks be uncorrelated with changes in (inverse) housing productivity, Δc_{it}, which measures how efficiently developers provide housing density. One potential concern with Assumption A-2 is through labor reallocation: If labor demand shocks alter the pool of workers available for construction, there could be cause for concern. Another is whether local zoning responds to local labor demand shocks. An alternative version drops the local labor demand shock in i:

$$\Delta z_{it}^{HD,a}(\rho) = \sum_{s \neq i} \frac{e^{-\rho\delta_{is}} 1_{\check{n}_{is}>0} \Delta z_{st}}{\sum_{s \neq i} e^{-\rho\delta_{is}} 1_{\check{n}_{is}>0}}$$

This spatially weighted labor demand shock identifies housing supply if $\mathbb{E}[\Delta z_{it}^{HD,a}(\rho) \times \Delta c_{it}] = 0, \ \forall i$, which can rewritten to be more easily parsed as a function of the labor demand shock itself:

$$k_{ij}\mathbb{E}[\Delta z_{jt} \times \Delta c_{it}] = 0, \ \forall \ i \neq j \tag{M-2a}$$

where $k_{ij} = e^{-\rho\delta_{ij}} 1_{\check{n}_{ij}>0}$ is the weight. Condition M-2a requires productivity shocks in a location be uncorrelated with innovations in nearby housing efficiency. Labor demand shocks have been used to estimate the aggregate housing supply elasticity (Diamond 2016; Guerrieri, Hartley, and Hurst 2013; Saiz 2010), but the use of spatially heterogeneous labor demand shocks at the census tract level to identify a local elasticity is novel.

Results and discussion

Implementation of the moment condition M-2a requires choosing the spatial decay parameter, $\rho > 0$, that governs how the labor demand shocks propagate across space. I experiment with different values in $\ln(\rho) \in [-10, -2]$.[32] Labor demand shocks should propagate through the econ-

32. The value of ρ impacts efficiency but not identification for $\rho \in (0, \infty)$. As $\rho \to 0$, the spatial correlation of the shocks increases. In the limit, there is no variation in the instrument, and the system is not identified. On the other hand, as $\rho \to \infty$, shocks do not influence activity elsewhere (autarky), and the system is not identified.

omy following the same decay as commuting, as these shocks will affect nearby markets only to the extent that workers are willing to commute to and from those markets. This implies $\rho = \epsilon\kappa$, and suggests using values of $\ln(\rho) \in [-7.5, -3.5]$ based on gravity estimates.[33] I report results for $\ln(\rho) = -5.5$.

Estimates of the inverse housing supply elasticity (Equation 16) appear in Table 5. Results are estimated in differences using the employment instrument. These estimates imply housing supply elasticities of about 0.45 when no adjustment is made for income-driven variation in quantity (Columns 1 and 2), and about 0.60 when income can influence housing quantity (Columns 3-6). Including available residential land decreases these estimates.[34] Column 2, 4, and 6 exclude the own tract labor demand shock when aggregating the instrument; this permits local housing productivity to covary with the local labor demand shock. Estimates are similar. All results suggest that local, tract-level housing provision is inelastic in the Los Angeles region from 1990 to 2000. Saiz (2010) finds the median long-run inverse housing supply elasticity among major U.S. metropolitan areas to be about 1.75; his estimate for the Los Angeles area is 0.63. My estimates similarly point to limited medium-run scope for adjustment in local housing stock. This matches anecdotal and empirical evidence on the highly regulated California housing market (Quigley and Raphael 2005).

6.4 Other parameters and robustness to agglomeration and endogenous land use

I show in the Appendix how further interactions of labor demand shocks with urban commuting geography can identify the remaining model parameters, the elasticities of labor and housing demand. However, the underlying moment conditions are potentially less plausible than M-1a and M-2a. Therefore, I take parameters from microdata or the literature. I use $\zeta = 0.65$, implying that the household expenditure share on housing is $1 - \zeta = 0.35$ and that the elasticity of housing demand is $-\epsilon(1-\zeta) = -0.64$. I assume that labors share in production is $\alpha = 0.68$. These values of these parameters matter little for welfare results, but are important for correctly recovering local fundamentals. Interestingly, when I identify these parameters using further interactions of labor demand shocks with geography, estimates come reasonably close.

I abstract from agglomerative forces and endogenous land use determination (agglomeration was included in prior versions and had little quantitative effect, and land use change is very rare in the data). I summarize here how relaxing this alters identification; details are in the Appendix. Identification of ϵ and ψ is still possible when agglomeration influences productivity and

33. I recover gravity-based estimates of the cofficient of flows on (linear) travel time ranging between 0 and -0.03, depending on specification and the inclusion of pair-fixed effects.

34. As shown in the Appendix, $h_i = n_i^H + \ln(\bar{W}_i^r) - \ln(L_i^H) - q_i + \text{constant}$, where n_i^H is residential population and $\ln(\bar{W}_i^r)$ is average residential wage (the housing price term drops to the left-hand side). Columns 1 and 2 of Table 5 assume $h_i = n_i^H - \ln(L_i^H)$, and Columns 5 and 6 assume $h_i = n_i^H + \ln(\bar{W}_i^r) - \ln(L_i^H)$. Columns 3 and 4 interrogate the assumption of Cobb-Douglas housing production by parameterizing $\psi h_i = \psi_0(n_i^H + \ln(\bar{W}_i^r)) + \psi_1 \ln(L_i^H)$. If correct, ψ_k should be equal in magnitude and opposite in sign. The coefficients are not statistically different in absolute value.

residential amenities, though the demand elasticities can no longer be identified without ex ante knowledge of the agglomeration or land use parameters. Agglomeration tends to be highly path dependent, therefore fixed effects $\bar{a}_j$ and $\bar{b}_i$ mostly control for these forces (Davis and Weinstein 2008). The model phrases labor demand and housing supply in terms of employment and housing density, respectively. Land use is relatively fixed and primarily captured by the tract fixed effects $\bar{a}_i$ and $\bar{c}_i$. However, I observe measures of land use (zoning in 1990 and 2001) for housing and production that let me measure density and examine changes in land use (they are small). Moment condition M-2 requires productivity shocks to be orthogonal to changes in a location's ability to provide employment density and housing density; it is density, rather than levels, that matters. Further, condition M-2a removes concerns about local productivity shocks driving land use changes in the same location.

7 Non-commuting effects of transit

Given parameters $\{\epsilon, \psi, \alpha, \zeta\}$ and data on workplace wages, residential housing prices, and commuting, the model delivers straightforward expressions to recover local economic fundamentals. These economic fundamentals represent economic characteristics of a place that exist outside of a market equilibrium. In combination with market forces, fundamentals determine equilibrium prices and the distribution of people. The fundamentals contain information about local productivity, housing supply, and transportation networks, and can be used to study how policy interventions generally shift supply and demand.

Consider a local intervention, T. In general, the intervention could impact any local fundamental. The following econometric framework permits estimating the effect of the intervention on local fundamentals:

$$\hat{\mathbf{Y}}_{it} = \boldsymbol{\lambda} T_{it} + \boldsymbol{\varsigma}_i + \boldsymbol{\varepsilon}_{it} \tag{17}$$

where $\boldsymbol{\lambda} = \{\lambda^A, \lambda^B, \lambda^C, \lambda^E\}$ are the effects to be estimated and $\hat{\mathbf{Y}} = \{\widehat{\mathbf{a}}, \widehat{\mathbf{b}}, \widehat{\mathbf{c}}, \widehat{\mathbf{e}}\}$ contains the four non-commuting fundamentals (I show in Section 8 how the commuting effect and non-commuting effects are combined for welfare calculations). Standard econometric techniques (e.g., difference-in-difference, instrumental variables) can then be performed on the above system. While the full sample should be used to estimate the structural elasticities, the effects of interventions can be estimated using a restricted sample if needed to overcome selection bias.[35]

35. Context or theory may dictate additional restrictions, i.e., some λ may be zero. Interestingly, when some λ can be assumed equal to zero, and others are non-zero, treatment can be used as an instrument to identify some or all of the structural elasticities. The economic geography literature typically assumes $\lambda^D \neq 0$ and $\boldsymbol{\lambda} = \{\lambda^A, \lambda^B, \lambda^C, \lambda^E\} = 0$.

7.1 The effects of the LA Metro on non-commuting fundamentals

I now test whether transit shifts non-commuting fundamentals using Equation (17). For single tract (rather than tract pair) analysis, I define transit proximity as

$$\text{Proximity}_i^{500m} = \frac{\max\{0, 500m - \min_k\{\text{dist}_i(\text{MetroStation}_k)\}\}}{500m} \in [0,1]$$

where the minimization is across stations. This normalizes proximity so that it equals one when a tract contains a station, and zero if a tract is more than 500m from a station. As before, I use the historical subway plan and location of PER lines to define control groups. I cannot define fundamentals prior to 1990, but as before, I can compare pre-trends on observed market outcomes using NCDB data. Appendix Table F1 shows that there do not appear to be consistent differences in treated and control locations among modeled prices and quantities across specifications. I include the 1990 levels of sociodemographic variables (income, education, and manufacturing employment) in some specifications to check robustness to differential trends.

While Section 4 showed strong evidence that transit increases commuting between connected locations, there is little evidence it affects other margins: Transit does not consistently shift local fundamentals. Estimates of the effect of transit on productivity, residential amenities, housing productivity, and workplace amenities are shown in Table 6.[36] Transit has little non-commuting effect after conditioning on regional trends and changes in highway structure. Taken together, these results indicate that transit is not generating large-scale non-commuting benefits or costs within the immediate proximity of stations. An implication is that hedonic estimates of the effect of transit reflect a commuting benefit rather than other related neighborhood amenities.

Surprisingly, there is no non-commuting effect of LA Metro Rail on productivity (λ^A, Panel A of Table 6) or residential amenities (λ^B, Panel B).[37] These margins have been the subject of considerable research. Chatman and Noland (2014) and Duranton and Turner (2012) find evidence that transit increases productivity and leads to employment growth using metropolitan-level analysis, while Kahn (2007) and Billings (2011) show some gentrifying effects of transit and that transit can anchor local economic development. The experience of Los Angeles does not mirror such results, and are consistent with Schuetz (2015), who shows that retail (a consumption amenity) does not increase near new rail transit stations in California between 1992 and 2009.

An important caveat is that these results only apply to LA Metro between 1990 and 2000; I cannot extend the non-commuting analysis to more recent years (I do expand the commuting analysis in Section 9). The network was limited in size and connectivity at this time. As the transit network

36. Estimates of these effects using fundamentals recovered with other elasticities are in the Appendix.

37. Estimates of λ^C and λ^E are worth discussing, though it is highly unlikely that transit itself can shift either margin. An effect on housing efficiency could occur if land use regulation was tied to transit accessibility (as recent proposals in California have advocated). Panel D suggests transit could be correlated with non-commuting decreases in labor supply, though these are noisy. It is difficult to imagine how such a change could be due to transit, however (e.g., reduced parking).

has expanded, it has become more valuable in terms of transportation connectivity. Responses that depend on scale, or are slower to respond, could potentially manifest in recent years. LA Metro Rail's primary effect between 1990-2000 is to expand commuter connectivity in Los Angeles: The city can accommodate more people with transit.

7.2 Robustness checks: Sorting and land use

Transit users may differ from those who do not use transit (Glaeser, Kahn, and Rappaport 2008; LeRoy and Sonstelie 1983). If so, transit could induce equilibrium sorting. While the data limit the explicit addition of heterogeneity to the model, I find no evidence of differential trends in median household income between treated and control census tracts (see Table F8; results are small and insignificant). Figure 4 shows the relationship between transit and rail usage by income centiles in 1990 and 2000, and reveals no relationship between income and rail usage.

The model assumes predetermined land use. While identification of the structural elasticities and transit effects are robust to this, counterfactual simulation may or may not be. I use SCAG zoning maps to test this channel and find little evidence of association between land use change and treatment. Table F8 indicates the change in residential land use is very small and statistically insignificant across richer fixed effects specifications. This is unsurprising given the relatively fixed nature of land use in urban settings due to the slow depreciation of buildings and very strict zoning and land use regulation in the LA region.

8 Welfare calculations and additional quantitative exercises

To estimate counterfactuals, I employ the succinct hat notation of Dekle, Eaton, and Kortum (2008), letting $\hat{X}_{it} = X'_{it}/X_{it}$ represent the relative change of an observed or estimated variable X to its counterfactual value X'. This approach avoids using levels of unobserved fundamentals. Results are easily interpretable and given as a ratio to the observed price or population level. Furthermore, after solving the model in terms of updated equilibrium prices and populations, estimation of the counterfactual proceeds easily via an iterative algorithm that quickly finds a fixed point representing a counterfactual equilibrium. I lay out the algorithm in the Appendix.

I estimate counterfactual values of $\hat{\mathbf{W}}, \hat{\boldsymbol{\pi}}, \hat{\mathbf{Q}}$ (and sometimes $\hat{\bar{N}}$) relative to observed data in 2000 under various combinations of the estimated structural elasticities $\{\alpha, \epsilon, \zeta, \psi\}$. Using estimates of the fundamental effects of transit from the preceding section, I define alternative scenarios by adjusting fundamentals so $\hat{X}_i = 1 - \lambda^X T_i$, for $X \in \{A, B, C, D, E\}$. The assumption of an open or closed city plays an important role. In a closed city, total population does not adjust. This means that there are real utility gains; these gains are equalized across the city through general equilibrium movements in prices. The model delivers a simple expression for welfare changes as a function of changes in local fundamentals and prices—a hat-notation variant of Equation (6):

$$\%\Delta \text{ Welfare } \approx \ln \hat{\bar{U}} = \frac{1}{\epsilon} \ln \left(\frac{\hat{B}_i \hat{E}_j \hat{D}_{ij} \hat{W}_j^{*\epsilon} \hat{Q}_i^{*-\epsilon(1-\zeta)}}{\hat{\pi}_{ij}^{*}} \right) \tag{18}$$

for any pair ij, where $\hat{X}^*$ indicates the equilibrium value of X in the counterfactual under autarky (that is, fixing $\hat{\bar{N}} = 1$). Because utility is homogeneous of degree one in wage, a proportional change in utility translates to an equivalent proportional change in wage. To convert this to levels, I multiply the proportional change in utility by the average annual wage ($31,563) and aggregate population of workers (6.73 million) in 2000.

Instead, if the city is open, aggregate population adjusts so that the expected utility in the city is equivalent to $\bar{U}$. Thus aggregate welfare for incumbent residents is unchanged: No spatial arbitrage means that the expected utility of city residence is $\bar{U}$ both before and after the change in fundamentals, so I instead report changes in total population. This statistic captures the change in the population the city can accommodate under transit with no change to utility.

Annualized costs combine two elements: (i) operating subsidies and (ii) annualized capital expenditures. The annual operating subsidy for the rail portion of LA Metro's operations for FY 2001-2002 is about $162 million (2016 dollars). Total system cost for lines and stations completed by 1999 is $8.7 billion (2016 dollars). Annualizing this expense involves an element of taste. LA Metro's borrowing terms at the time were about 6%, so the annual payment for a 30-year loan is roughly $635 million. However, subways often last for a very long time once built. It may be appropriate to use a much lower social discount rate (see Weitzman 1998). Assuming a social discount rate of 2.5% and an infinite horizon, capital expenditures are equivalent to $218 million per year. Combining with the operating subsidy yields an annualized cost between $380 million and $797 million per year (details in the Appendix).[38]

Welfare effects by 2000

Table 7 reports the changes in aggregate welfare and population due to LA Metro in percentage and dollar terms. Estimates of the effect on commuting use results from Table 2, Column 5, the Subway Plan (All) sample. Columns 1 and 3 exclude travel benefits for automobile users, Columns 2 and 4 include them. All columns use $\epsilon = 1.83$, $\alpha = 0.68$, and $\zeta = 0.65$; Columns 1 and 2 use $\psi = 1.69$, and Columns 3 and 4 use $\psi = 2.29$.

Column 1 indicates an annual benefit of $109 million, an increase of 0.05% relative to baseline. In an open economy, the employed population of the Los Angeles region is 0.11% higher with

38. Alternatively, I estimate the deadweight loss to the region generated by LA Metro Rail's enabling sales tax. I use the portion of Proposition A (1980) and Proposition C (1990) exclusively assigned to rail ($0.004/dollar). I assume full incidence on the composite consumption good for LA County residents on a base tax of $0.06/dollar. Deadweight loss accounting for migration is -$298 million annually, a bit smaller but roughly similar to other cost estimates.

Metro Rail. This measure ignores any congestion benefits. If, instead, congestion savings from Section 4.3 are included, the closed economy benefit increases to about \$146 million per year, or about 0.07%. Open economy employment would be about 0.15% higher in this case (roughly 10,000 people). Closed economy results are virtually unchanged with less elastic housing supply, while open economy results are slightly smaller. Accounting for agglomeration leads to little quantitative change.[39]

A general conclusion across all specifications is that the commuting benefit of rail transit in Los Angeles does not exceed its cost by the year 2000. Regardless of the discount rate, annual benefits are almost equal to the operating subsidy (about \$160 million) in some specifications. However, these commuting benefits cannot cover the capital expenses, except at very low discount rates. Combining both expenses, costs clearly outweigh benefits by the end of the first decade of LA Metro Rail operation.

Discussion and other margins

The negative cost-benefit calculations are subject to particular characteristics of LA Metro Rail. It is important to understand these characteristics when drawing broader conclusions. I focus on two: the connectivity provided by rail transit and zoning. Early system planners considered political factors as much, or more than, the utility of route placement. While this may have been necessary to ensure feasibility and political support, rail provided a substitute for few commuters. Only 1%-3% of the 1990 population of Los Angeles County lived and worked in tracts near rail stations by 2000 (Table 1). Figure 5 shows that only about 15% of the tracts with the highest ex ante commuting that could plausibly have received transit (those tracts in the Subway Plan (All) sample) were served by 2000. Connecting denser corridors (such as Wilshire) would have generated greater gains.

Land use regulations also inhibited the ability of locations receiving stations to adjust building stock. Table F8 reveals that essentially no land was converted to residential use near LA Metro Rail stations. Furthermore, Proposition U, passed in 1986, halved allowable density throughout much of Los Angeles just before LA Metro Rail was brought online. Such legislation combined with political constraints meant the "coordinated land use and rail planning . . . died a gory death" (Elkind 2014, p. 71).

Finally, there are margins to which this analysis does not speak. City-wide effects are not captured by this approach. Decreased air pollution may provide an additional benefit; a gener-

39. There are two margins to consider: At the metropolitcan level, consider a simple agglomerative force that increases productivity by 5% everywhere for each doubling of city population (Ciccone and Hall 1996). Welfare is homogeneous of degree one in wage, and wage is proportional to productivity, so the additional welfare effect is capped at $\ln(1+0.05\times 0.0015) \approx 0.0075\%$. Local agglomeration is the other margin; prior versions included a local measure of agglomeration across a grid of parameterizations and found quantitatively insignificant effects. This is because LA Metro Rail does not substantially increase employment at any single destination, despite increasing commuting along specific connected origin-destination pairs.

ous estimate using parameters from Gendron-Carrier et al. (2018) and Los Angeles' mid-1990s birthrate suggests an additional gain of up to $180 million annually, though such benefits may be transitory.[40] Nor can I directly speak to benefits resulting in better transit provision for non-commuting trips, though this margin should show up as a residential amenity (which I do not find). This framework does not capture the benefits for non-workers. Such effects are particularly important for equity concerns and unfortunately understudied.

9 Longer-run commuting effects

Because LA Metro Rail was still relatively new in 2000 (some stations had only been built in the previous five years), I extend the commuting analysis to search for additional effects of transit on commuting flows in more recent years. I use data from the 2002 and 2015 LEHD Origin-Destination Employment Statistics (LODES). Because LA Metro Rail expanded during this period, I estimate a variant of Equation (2) on the LODES panel with two different effects: (i) *New Transit* for the effect of new stations (built after 2002) on bilateral flows, and (ii) *Existing Transit* for the additional increase between tracts connected by stations built earlier (between 1990 and 2002). I retain the prior distance bins.

Results (shown in Table 8) indicate that new transit connections increase commuting by 10%-13% between tract pairs that both contained stations by 2015. For tract pairs slightly farther away, the increase is 5%-8%, and is insignificant using the PER sample. While these effects are substantial, they are smaller than the effects of connections between 1990 and 2000. This is likely because many of the stations built after 2002 connect locations that are more suburban and less transit-oriented. Tract pairs that had been previously connected by transit (before 2000) experience additional commuting growth by 2015: Pairs both containing a station show another 8%-11% increase in commuting, and tract pairs a bit further away show an additional 5%-9% in commuting. This is evidence that (i) aggregate commuting flows take decades to adjust to new transit modes (i.e., *habituation*), and/or (ii) there are increasing returns in transit network size.

A significant omission of the 1990-2000 welfare analysis in Section 8 is the exclusion of these later benefits. However, the results in Table 8 can be included in longer run welfare analysis if I assume there are no non-commuting benefits post 2000 (as between 1990 and 2000). There are then two cases to consider: (1) If increased commuting is due to habituation, the full commuting increase between previously connected stations is attributable to early system construction; (2) if, however, there are increasing returns in network size, increased commuting between existing stations is due to new stations and lines. As a lower bound, the additional benefit is zero. This allows comparing outcomes using the same capital and operating cost bases.

40. The time frame on these benefits is uncertain, as Gendron-Carrier et al. (2018) are unable to measure pollution responses beyond six to eight years after the opening of a metro system.

Under habituation, I combine the effects from Table 2 and the *Existing Station* effects from Table 8 and simulate the new outcome. Accounting for these additional effects, the closed economy benefit is $216 million annually, or an increase of about 0.10% (using parameters from Column 2 of Table 7). In an open economy, population is about 0.22% higher because of Metro Rail. The benefit, while substantial, only exceeds operational subsidies and capital costs if the social discount rate is very low (about 0.6%).

10 Conclusion

This paper develops and estimates an equilibrium model of a city wherein costly commuting connects housing and labor markets, and uses this model to estimate the welfare impacts of Los Angeles Metro Rail. The model is sufficiently parsimonious to permit transparent identification and estimation of all parameters, yet better reflects the observed spatial distribution of economic activity than commonly used market access approaches. The elasticity of labor supply plays a key role governing homogeneity in location preference. A small value indicates agents are relatively unwilling to relocate and are not very responsive to changes in local conditions or policies. Conversely, it also implies that observed responses to transit correspond to significant utility gains. Estimates of the remaining elasticities are in line with previous studies, and support the view that Southern California has a constrained housing supply. I also demonstrate that common implementations of economic geography models within cities are not consistent with commuting flow and workplace wage data.

I provide new insights into how transit influences city structure by isolating the commuting benefit of transit from other margins. LA Metro Rail increases commuting between the census tracts nearest to stations by 15% in the first decade after construction, relative to control groups selected by proposed and historical transit routes. Nearby locations also experience a more modest increase of about 10%. There is some evidence that Metro Rail has a small medium to long run effect on congestion, reducing travel times in nearby areas by up to 3%. There is little support of effects through other channels (such as non-commuting amenities).

Welfare estimates point to a range of positive annual benefits of the system from $109 million to $146 million by 2000. These welfare benefits are smaller than the operational and capital costs of LA Metro's light rail and subway lines. I also provide evidence of dynamic effects due to increasing returns or habituation. If these effects are due to slow habituation, the annual benefits of LA Metro Rail's network are about $216 million. This benefit is greater than the operational subsidy, but only approaches the cost of capital under very low social discount rates. While these welfare estimates leave out some other benefits of transit (such as benefits for non-workers), results warrant a note of caution to cities expecting rail investment to lead to large increases in worker welfare within 10 to 25 years.

References

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales. 2019. "Shift-Share Designs: Theory and Inference." *The Quarterly Journal of Economics* 134 (4): 1949–2010.

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.

Albouy, David, and Bryan Stuart. 2014. "Urban population and amenities: The neoclassical model of location." *NBER Working Paper*, no. w19919.

Allen, Treb, Costas Arkolakis, and Xiangliang Li. 2014. "On the existence and uniqueness of trade equilibria."

———. 2015. "Optimal City Structure."

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

Baier, Scott L, and Jeffrey H Bergstrand. 2007. "Do free trade agreements actually increase members' international trade?" *Journal of International Economics* 71 (1): 72–95.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* Books from Upjohn Press. W.E. Upjohn Institute for Employment 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.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. "Migration and hedonic valuation: The case of air quality." *Journal of Environmental Economics and Management* 58 (1): 1–14.

Bayer, Patrick, and Christopher Timmins. 2005. "On the equilibrium properties of locational sorting models." *Journal of Urban Economics* 57 (3): 462–477.

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

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

Bowes, David R, and Keith R Ihlanfeldt. 2001. "Identifying the impacts of rail transit stations on residential property values." *Journal of Urban Economics* 50 (1): 1–25.

Brooks, Leah, and Byron Lutz. 2016. "Vestiges of Transit: Urban Persistence at a Micro Scale."

Chatman, Daniel G, and Robert B Noland. 2014. “Transit service, physical agglomeration and productivity in US metropolitan areas.” *Urban Studies* 51 (5): 917–937.

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.

Ciccone, Antonio, and Robert E Hall. 1996. “Productivity and the Density of Economic Activity.” *American Economic Review* 86 (1): 54–70.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2012. “The costs of agglomeration: House and land prices in French cities.”

Davis, Donald R, and David E Weinstein. 2008. “A search for multiple equilibria in urban industrial structure.” *Journal of Regional Science* 48 (1): 29–65.

Davis, Morris A, and François Ortalo-Magné. 2011. “Household expenditures, wages, rents.” *Review of Economic Dynamics* 14 (2): 248–261.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2008. “Global rebalancing with gravity: Measuring the burden of adjustment.” *NBER Working Paper,* no. w13846.

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.

Donaldson, Dave, and Richard Hornbeck. 2016. “Railroads and American economic growth: A “market access” approach.” *Quarterly Journal of Economics* 131 (2): 799–858.

Downs, Anthony. 1962. “The law of peak-hour expressway congestion.” *Traffic Quarterly* 16 (3).

Duranton, Gilles, and Matthew A Turner. 2011. “The fundamental law of road congestion: Evidence from US cities.” *American Economic Review* 101 (6): 2616–2652.

———. 2012. “Urban growth and transportation.” *Review of Economic Studies* 79 (4): 1407–1440.

Elkind, Ethan N. 2014. *Railtown: The fight for the Los Angeles metro rail and the future of the city.* University of California Press.

Epple, Dennis, Brett Gordon, and Holger Sieg. 2010. “A new approach to estimating the production function for housing.” *American Economic Review* 100 (3): 905–924.

Epple, Dennis, and Holger Sieg. 1999. “Estimating Equilibrium Models of Local Jurisdictions.” *Journal of Political Economy* 107 (4): 645–681.

Falch, Torberg. 2010. “The elasticity of labor supply at the establishment level.” *Journal of Labor Economics* 28 (2): 237–266.

Fogelson, R.M. 1967. *The Fragmented Metropolis: Los Angeles, 1850-1930.* University of California Press.

Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew A Turner. 2018. “Subways and urban air pollution.” *NBER Working Paper* w24183.

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

Glaeser, Edward L, et al. 2008. *Cities, agglomeration, and spatial equilibrium.* Oxford University Press.

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.

Glaeser, Edward L, Jed Kolko, and Albert Saiz. 2001. "Consumer city." *Journal of Economic Geography* 1 (1): 27–50.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2018. "Bartik instruments: What, when, why, and how." *NBER Working Paper,* no. w24408.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst. 2013. "Endogenous gentrification and housing price dynamics." *Journal of Public Economics* 100:45–60.

Hsu, Wen-Tai, and Hongliang Zhang. 2014. "The fundamental law of highway congestion revisited: Evidence from national expressways in Japan." *Journal of Urban Economics* 81:65–76.

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.

Kahn, Matthew E, Ryan Vaughn, and Jonathan Zasloff. 2010. "The housing market effects of discrete land use regulations: Evidence from the California coastal boundary zone." *Journal of Housing Economics* 19 (4): 269–279.

Kelker, De Leuw & Company. 1925. *Report and Recommendations on a Comprehensive Rapid Transit Plan for the City and County of Los Angeles.* Technical report. Chicago.

Kreindler, Gabriel, and Yuhei Miyauchi. 2017. "Billions of Calls Away from Home: Measuring Commuting and Productivity inside Cities with Cell Phone Records."

Larch, Mario, Joschka Wanner, Yoto V Yotov, and Thomas Zylkin. 2019. "Currency Unions and Trade: A PPML Re-assessment with High-dimensional Fixed Effects." *Oxford Bulletin of Economics and Statistics* 81 (3): 487–510.

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.

Logan, John R, Zengwang Xu, and Brian J Stults. 2014. "Interpolating U.S. decennial census tract data from as early as 1970 to 2010: A longitudinal tract database." *Professional Geographer* 66 (3): 412–420.

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.

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg. 2018. "Commuting, Migration, and Local Employment Elasticities." *American Economic Review* 108 (12): 3855–90.

Quigley, John M, and Steven Raphael. 2005. "Regulation and the high cost of housing in California." *American Economic Review* 95 (2): 323–328.

Redding, Stephen J, and Matthew A Turner. 2015. "Transportation costs and the spatial organization of economic activity," 5th ed., edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange, 1339–1398. Handbook of Regional and Urban Economics. Elsevier.

Redfearn, Christian L. 2009. "How informative are average effects? Hedonic regression and amenity capitalization in complex urban housing markets." *Regional Science and Urban Economics* 39 (3): 297–306.

Roback, Jennifer. 1982. "Wages, rents, and the quality of life." *Journal of Political Economy* 90 (6): 1257–1278.

Rosen, Sherwin. 1979. "Wage-based indexes of urban quality of life." *Current Issues in Urban Economics:* 74–104.

Saiz, Albert. 2010. "The geographic determinants of housing supply." *Quarterly Journal of Economics* 125 (3): 1253–1296.

Schrank, David, Bill Eisele, Tim Lomax, and Jim Bak. 2015. *Urban Mobility Scorecard.* Technical report. Texas A&M University.

Schuetz, Jenny. 2015. "Do rail transit stations encourage neighbourhood retail activity?" *Urban Studies* 52 (14): 2699–2723.

Schuetz, Jenny, Genevieve Giuliano, and Eun Jin Shin. 2018. "Does zoning help or hinder transit-oriented (re)development?" *Urban Studies* 55 (8): 1672–1689.

Severen, Christopher, and Andrew J Plantinga. 2018. "Land-use regulations, property values, and rents: Decomposing the effects of the California Coastal Act." *Journal of Urban Economics* 107:65–78.

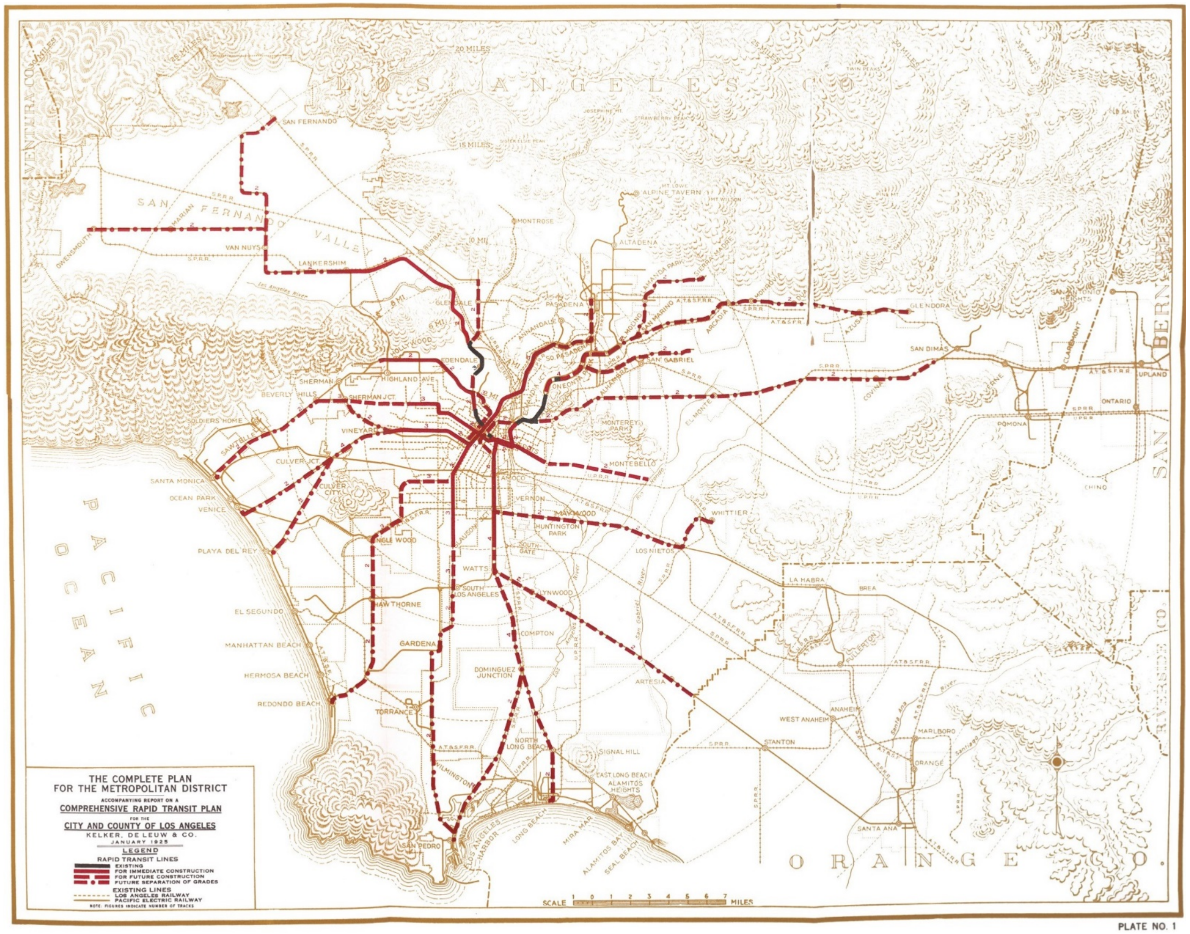
Suárez Serrato, Juan Carlos, and Owen Zidar. 2016. "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms." *American Economic Review* 106 (9): 2582–2624.

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

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

Weitzman, Martin L. 1998. "Why the far-distant future should be discounted at its lowest possible rate." *Journal of Environmental Economics and Management* 36 (3): 201–208.

Figure 1: Map of Proposed LA Metro Lines and PER Lines in Kelker, De Leuw & Co. (1925)



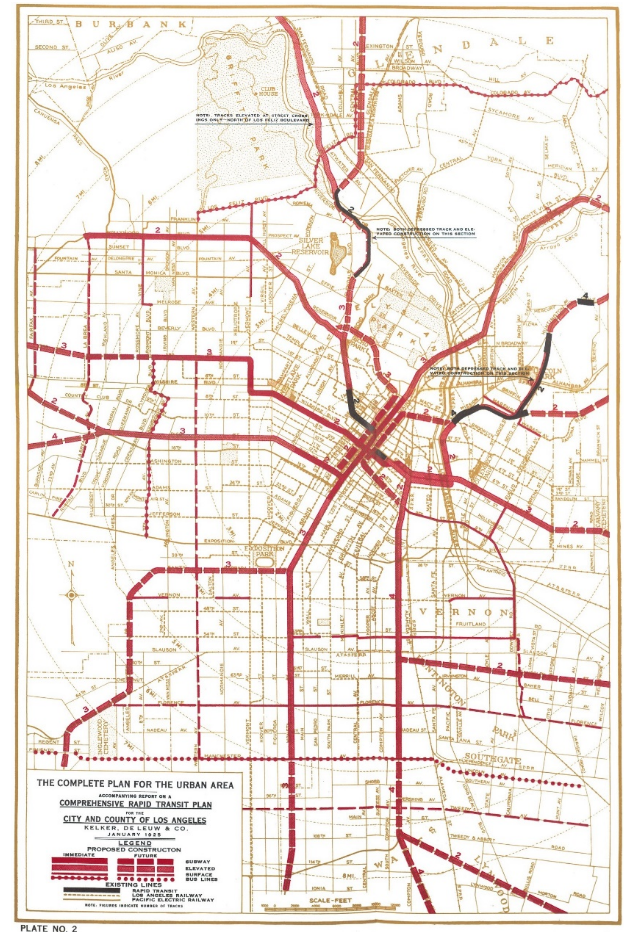
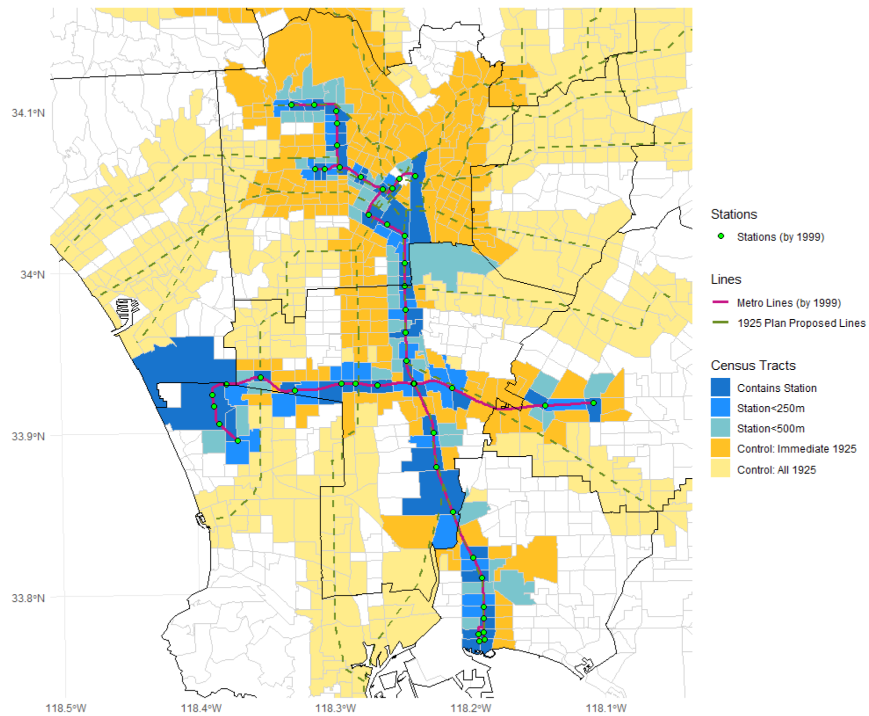


Figure 2: Map of LA Metro lines, stations, and the 1925 Plan and PER Lines

(a) 1925 Plan Sample



(b) PER Lines Sample

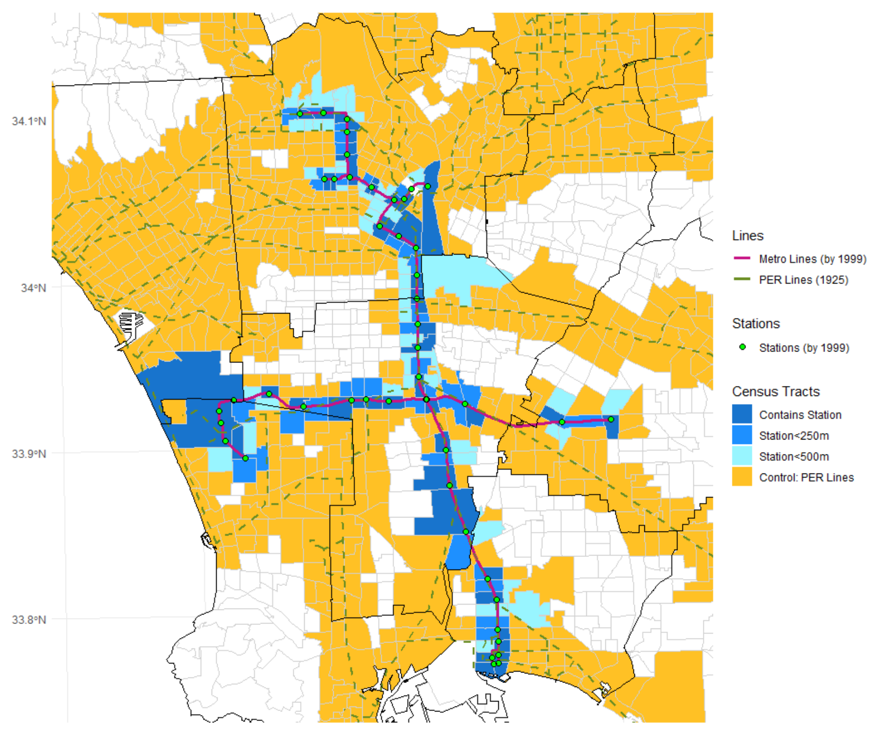


Figure 3: Does $\omega = \epsilon w$, and if not, what is it capturing?

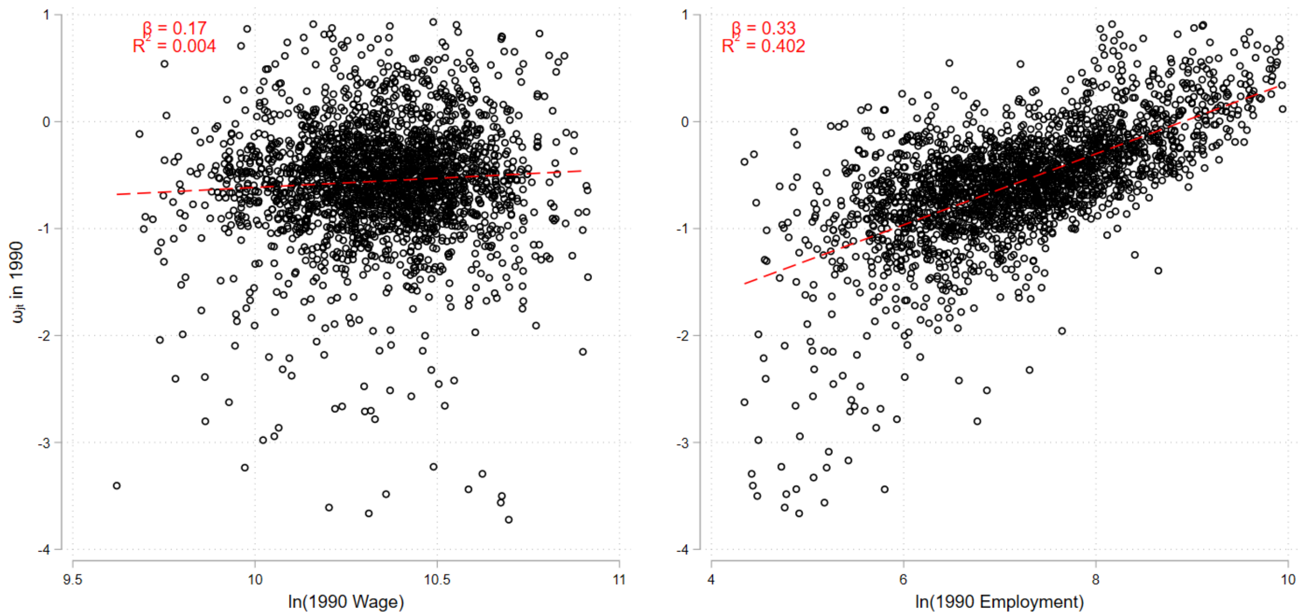


Figure 4: Take-up of LA Metro Rail for commuting does not vary by income, but overall take-up of transit (including bus) does.

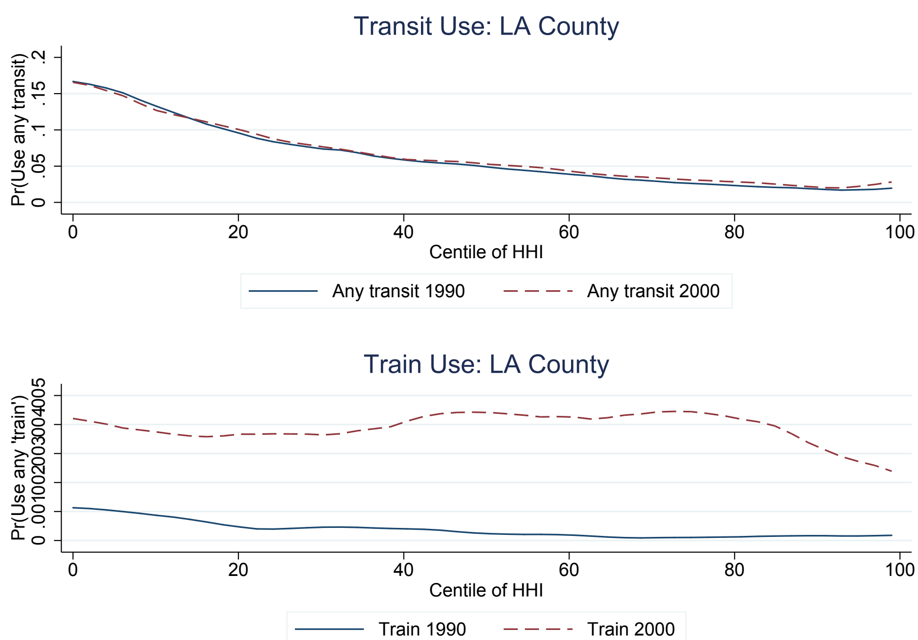


Figure 5: Tract pairs with higher ex-ante commuting flows are a bit more likely to receive LA Metro Rail by 2000, but many high commuting pairs do not.

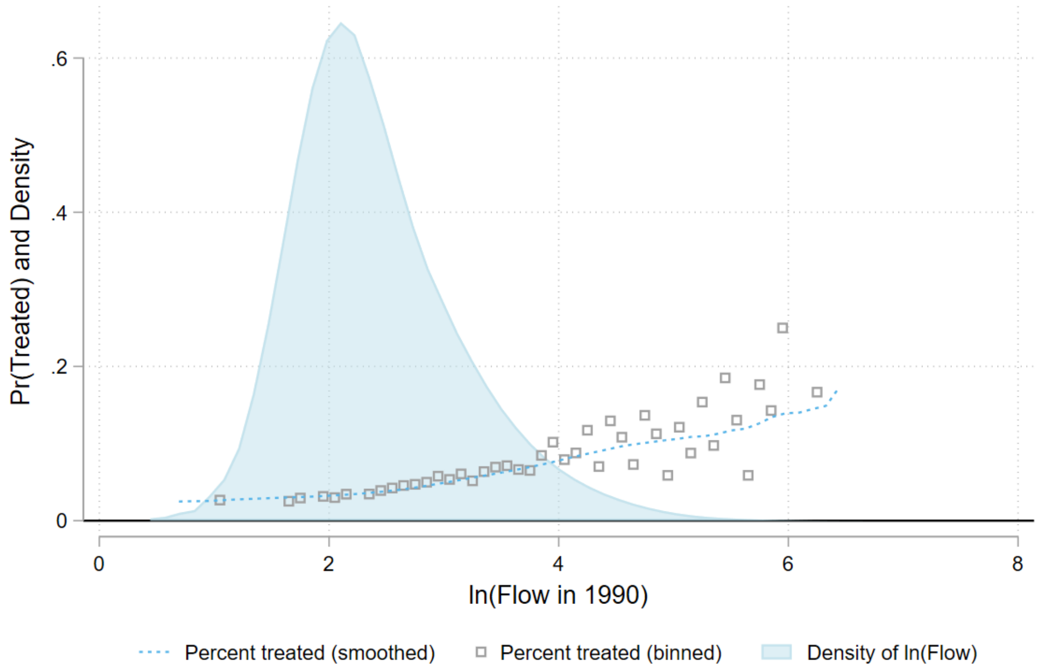


Table 1: Descriptive statistics on transportation in Los Angeles and station placement

	LA County		Full Sample	
	Centroid $< 500m$ (1)	Any $< 500m$ (2)	Centroid $< 500m$ (3)	Any $< 500m$ (4)
A. Pre-treatment tract characteristics (1990)				
% workers at POW tract that receive treatment	11.3%	19.5%	7.2%	12.3%
% workers at RES tract that receive treatment	2.7%	8.1%	1.6%	4.8%
% workers that receive transit connection RES-POW	0.6%	2.9%	0.4%	1.7%
% workers commuting via: Drive alone	71.8%		74.5%	
% workers commuting via: Carpool	15.8%		15.8%	
% workers commuting via: Bus	6.9%		4.6%	
B. Commuting characteristics				
Commute time (minutes, 1990)	26.3 [16.8]			
Commute time (minutes, 2000)	28.0 [18.3]			

Data from Census micro records (from IPUMS) and 1990 CTPP. LA County restricts analysis to workers both living and residing in Los Angeles county, while the full sample includes all five counties in the main sample. Brackets indicate standard deviation. Commute times are weighted by flows.

Table 2: Effect of Transit on Commuting Flows (by 2000) - Linear

	$n_{ijt} = \ln(N_{ijt})$					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway Plan (Immediate) Sample						
O & D contain station	0.106*	0.142*	0.153*	0.157*	0.149*	0.154*
	(0.052)	(0.059)	(0.060)	(0.061)	(0.061)	(0.066)
O & D <250m from station			0.127*	0.138*	0.128*	0.137*
			(0.063)	(0.064)	(0.065)	(0.069)
O & D <500m from station		0.058	0.017	0.016	0.012	0.018
		(0.048)	(0.053)	(0.053)	(0.053)	(0.055)
N	19238	19238	19238	19222	19222	18296
Subway Plan (All) Sample						
O & D contain station	0.127**	0.147**	0.152**	0.162**	0.146**	0.152**
	(0.044)	(0.044)	(0.044)	(0.046)	(0.044)	(0.047)
O & D <250m from station			0.115*	0.122*	0.101*	0.109*
			(0.049)	(0.050)	(0.051)	(0.053)
O & D <500m from station		0.054	0.018	0.023	0.013	0.018
		(0.035)	(0.044)	(0.042)	(0.042)	(0.043)
N	74046	74046	74046	74040	74040	71844
PER Sample						
O & D contain station	0.098*	0.116**	0.119**	0.129**	0.113*	0.084^{+}
	(0.042)	(0.043)	(0.043)	(0.045)	(0.044)	(0.046)
O & D <250m from station			0.104*	0.109*	0.088^{+}	0.037
			(0.049)	(0.050)	(0.051)	(0.050)
O & D <500m from station		0.054	0.025	0.030	0.019	-0.027
		(0.034)	(0.041)	(0.040)	(0.040)	(0.043)
N	99074	99074	99074	99054	99054	95382
Full Sample						
O & D contain station	0.102**	0.101**	0.112**	0.117**	0.101**	
	(0.038)	(0.037)	(0.038)	(0.040)	(0.038)	
O & D <250m from station			0.074	0.077^{+}	0.054	
			(0.046)	(0.046)	(0.047)	
O & D <500m from station		0.028	0.000	-0.003	-0.014	
		(0.031)	(0.037)	(0.036)	(0.036)	
N	291000	291000	291000	290580	290580	
Control Network	Loose	Loose	Loose	Loose	Loose	Tight
Tract Pair FE	Y	Y	Y	Y	Y	Y
POW-X-Yr FE	Y	Y	Y	Y	Y	Y
RES-X-Yr FE	Y	Y	Y	Y	Y	Y
Sbcty-X-Sbcty-X-Yr FE	-	-	-	Y	Y	Y
Highway Control	-	-	-	-	Y	Y

High-dimensional fixed effects estimates of λ^D. Treatment variables are mutually exclusive with others in each column. All estimates include tract of work-by-year, tract of residence-by-year, and tract pair fixed effects. Standard errors clustered by tract pair, tract of residence, and tract of work in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 3: Does transit decrease travel time?

	τ_{ijt}^{All}			$\ln(\tau_{ijt}^{\text{All}})$			τ_{ijt}^{Car}	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Within 2km of tracks	-1.277**	-1.243**	-0.748	-0.032*	-0.033*	-0.026	-1.417*	-0.766
	(0.402)	(0.426)	(0.481)	(0.013)	(0.014)	(0.016)	(0.631)	(0.719)
Within 4km of tracks	-0.305	-0.304	0.050	-0.006	-0.006	-0.002	0.150	0.293
	(0.364)	(0.364)	(0.398)	(0.012)	(0.012)	(0.013)	(0.556)	(0.612)
Control Network	All	All	All	All	All	All	All	All
Tract Pair FE	Y	Y	Y	Y	Y	Y	Y	Y
POW-X-Yr FE	Y	Y	Y	Y	Y	Y	Y	Y
RES-X-Yr FE	Y	Y	Y	Y	Y	Y	Y	Y
Near station control	-	Y	Y	-	Y	Y	-	-
Sbcty-X-Sbcty-X-Yr FE	-	-	Y	-	-	Y	-	Y
Highway Control	-	-	Y	-	-	Y	-	Y
N	311340	311340	310904	311314	311314	310878	96098	95884

High-dimensional fixed effects estimates of track proximity on driving time. Control network is 'loose' (see text). Treatment variables are mutually exclusive with others in each column. All estimates include tract of work-by-year, tract of residence-by-year, and tract pair fixed effects. Standard errors clustered by tract pair, tract of residence, and tract of work in parentheses: $^{+}\ p < 0.10$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$

Table 4: IV estimates of labor supply elasticity (ϵ)

	$\hat{\omega}_{jt}$		
	(1)	(2)	(3)
$\ln(W_{jt})$	0.498	1.846*	1.830*
	(0.411)	(0.835)	(0.783)
F-stat (KP)	15.277	16.883	17.328
$\hat{\omega}$ estimated:	Linear, Panel	PPML Yr-by-yr	PPML Panel
N	2354	2432	2433

Panel instrument variable (IV) estimates of regression of $\hat{\omega}_j t$ on w_{it}. Estimated in differences using wage instrument. KP refers to the Kleinbergen-Papp F-statistic. Variables are trimmed to exclude extreme values (see text). Column 1 uses a linear specification, columns 2 and 3 assume a Poisson model. Place of work-by-year fixed effects ($\hat{\omega}_{jt}$) estimated in the panel in columns 1 and 3 with ij fixed effects. Column 2 uses $\hat{\omega}_{jt}$ estimated year-by-year with network distance as a control. Standard errors clustered by tract in parentheses: $^{+}\ p < 0.10$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$

Table 5: IV estimates of inverse housing supply elasticity (ψ)

	$q_{it} = \ln(Q_{it})$					
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Density)	2.221**	2.292**				
	(0.706)	(0.738)				
ln(Hous. Consump.)			1.693**	1.610**		
			(0.483)	(0.442)		
ln(Res. Land)			-1.396^{+}	-1.318^{+}		
			(0.790)	(0.778)		
ln(Hous. Density)					1.814**	1.693**
					(0.648)	(0.504)
Housing Supply Elasticity ($1/\psi$)	0.450**	0.436**	0.591**	0.621**	0.551**	0.591**
	(0.143)	(0.140)	(0.169)	(0.170)	(0.197)	(0.176)
F-stat (KP)	14.830	14.234	12.944	14.218	8.138	11.887
Empl. instrument	All	Not i	All	Not i	All	Not i
N	4550	4548	4500	4498	4500	4498

Panel instrument variable (IV) estimates of regression of median house value on population, housing consumption, and residential land, using $\ln(\rho) = -5.5$ and employment IV. KP refers to the Kleinbergen-Papp F-statistic. Variables are trimmed to exclude extreme values (see text). Columns 2, 4, and 6 exclude own tract during instrument construction. Standard errors clustered by tract in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 6: Transit and non-commuting fundamentals (other effects of transit)

	$\Delta\hat{\mathbf{Y}}_{it}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Effect on productivity ΔA**,** $\alpha - 1 = -0.226$								
Proximity$_i^{500m} \times t$	-0.089**	0.009	-0.009	0.008	-0.034	0.006	-0.050^{+}	0.011
	(0.027)	(0.030)	(0.031)	(0.034)	(0.028)	(0.030)	(0.027)	(0.030)
N	4882	4858	780	776	1828	1826	2288	2284
B. Effect on residential amenity ΔB**,** $\epsilon(1-\zeta) = 0.662$								
Proximity$_i^{500m} \times t$	0.107**	-0.002	0.030	-0.042	0.070*	-0.007	0.076**	0.012
	(0.029)	(0.032)	(0.032)	(0.035)	(0.029)	(0.033)	(0.029)	(0.033)
N	4534	4518	712	710	1700	1700	2094	2092
C. Effect on inverse housing efficiency ΔC**,** $\psi = 1.693$								
Proximity$_i^{500m} \times t$	0.070^{+}	0.006	-0.096*	-0.044	0.024	-0.025	0.051	0.003
	(0.041)	(0.046)	(0.047)	(0.054)	(0.042)	(0.048)	(0.042)	(0.048)
N	4484	4476	694	692	1670	1670	2058	2056
D. Effect on workplace amenity ΔE**,** $\epsilon = 1.83$								
Proximity$_i^{500m} \times t$	-0.203**	-0.058	-0.092	-0.154*	-0.103^{+}	-0.103	-0.104^{+}	-0.115^{+}
	(0.058)	(0.062)	(0.066)	(0.073)	(0.060)	(0.066)	(0.059)	(0.062)
N	4866	4842	780	776	1830	1828	2286	2282
Sample	All	All	Sim	Sim	Sal	Sal	PER	PER
Tract FE	Y	Y	Y	Y	Y	Y	Y	Y
Sbcty-X-Yr FE	-	Y	-	Y	-	Y	-	Y
Controls	-	Y	-	Y	-	Y	-	Y

Results from 32 regressions of transit proximity on local fundamentals. All regressions include tract fixed effects. Samples: 'All' is the Full Sample, 'Sim' is Subway Plan (Immediate), 'Sal' is Subway Plan (All), and PER is the PER Sample. Controls include changes in highway proximity and 1990 levels of log household income, share of residents with at least a high school degree, and manufacturing employment. Standard errors clustered by tract in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 7: Welfare estimates in 2000 (in $2016)

	(1)	(2)	(3)	(4)
Parameters				
α	0.680	0.680	0.680	0.680
ϵ	1.830	1.830	1.830	1.830
ζ	0.650	0.650	0.650	0.650
ψ	1.693	1.693	2.290	2.290
$\epsilon\kappa$	-	-0.020	-	-0.020
Change in fundamentals				
λ^D, O & D contain station	0.146	0.146	0.146	0.146
λ^D, O & D <250m from station	0.101	0.101	0.101	0.101
λ^τ, O & D <2km from station	-	-0.033	-	-0.033
Closed Economy				
Annual Δ in welfare	0.051%	0.069%	0.051%	0.069%
(in millions of $2016)	108.9	145.7	108.9	145.6
Open Economy				
Population Δ	0.109%	0.146%	0.106%	0.141%
Op. subsidy + capital cost (6%, 30yr)		-$797 mil.		
Op. subsidy + capital cost (5%, 50yr)		-$641 mil.		
Op. subsidy + capital cost (5%, ∞)		-$597 mil.		
Op. subsidy + capital cost (2.5%, ∞)		-$380 mil.		
Operation subsidy only		-$162 mil.		

Op. subsidy refers to the annual operation subsidy. See text and appendices for details.

Table 8: Dynamic effects of transit on flows (2002-15), Linear

	$n_{ijt} = \ln(N_{ijt})$					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway Plan (Immediate) Sample, $N = 105794$						
New: O & D contain station	0.101**	0.103**	0.121**	0.123**	0.131**	0.133**
	(0.035)	(0.035)	(0.035)	(0.035)	(0.036)	(0.036)
New: O & D <250m from station	0.053+	0.053*	0.075**	0.076**	0.082**	0.083**
	(0.027)	(0.026)	(0.028)	(0.027)	(0.028)	(0.027)
New: O & D <500m from station	0.026	0.026	0.053*	0.054*	0.043+	0.045+
	(0.026)	(0.026)	(0.025)	(0.025)	(0.025)	(0.025)
Existing: O & D contain station					0.108**	0.112**
					(0.030)	(0.030)
Existing: O & D <250m from station					0.086**	0.091**
					(0.029)	(0.029)
Existing: O & D <500m from station			0.058**	0.061**	0.029	0.032
			(0.022)	(0.022)	(0.028)	(0.029)
Subway Plan (All) Sample, $N = 385290$						
New: O & D contain station	0.109**	0.102**	0.113**	0.106**	0.119**	0.112**
	(0.031)	(0.031)	(0.032)	(0.031)	(0.032)	(0.031)
New: O & D <250m from station	0.041+	0.036	0.050*	0.044+	0.052*	0.046*
	(0.023)	(0.023)	(0.024)	(0.023)	(0.024)	(0.023)
New: O & D <500m from station	0.019	0.016	0.034+	0.029	0.029	0.025
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Existing: O & D contain station					0.107**	0.098**
					(0.033)	(0.032)
Existing: O & D <250m from station					0.066+	0.061+
					(0.035)	(0.035)
Existing: O & D <500m from station			0.056*	0.049*	0.035	0.028
			(0.023)	(0.023)	(0.025)	(0.025)

continued...

Table 8 – continued from previous page

PER Sample, $N = 514110$						
New: O & D contain station	0.101**	0.097**	0.103**	0.100**	0.108**	0.105**
	(0.033)	(0.032)	(0.034)	(0.033)	(0.034)	(0.033)
New: O & D <250m from station	0.030	0.026	0.037	0.033	0.039	0.035
	(0.024)	(0.024)	(0.025)	(0.024)	(0.025)	(0.024)
New: O & D <500m from station	0.025	0.024	0.039+	0.038+	0.036+	0.035+
	(0.020)	(0.020)	(0.021)	(0.020)	(0.021)	(0.020)
Existing: O & D contain station					0.098**	0.099**
					(0.035)	(0.034)
Existing: O & D <250m from station					0.058+	0.059+
					(0.031)	(0.032)
Existing: O & D <500m from station			0.055*	0.056*	0.040	0.041
			(0.022)	(0.023)	(0.025)	(0.025)
Full Sample, $N = 1993198$						
New: O & D contain station	0.109**	0.092**	0.108**	0.092**	0.110**	0.094**
	(0.033)	(0.031)	(0.034)	(0.032)	(0.034)	(0.032)
New: O & D <250m from station	0.034	0.017	0.038	0.021	0.038	0.022
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
New: O & D <500m from station	0.022	0.008	0.032	0.019	0.031	0.017
	(0.022)	(0.020)	(0.022)	(0.020)	(0.023)	(0.021)
Existing: O & D contain station					0.091*	0.084*
					(0.038)	(0.036)
Existing: O & D <250m from station					0.049+	0.048+
					(0.027)	(0.029)
Existing: O & D <500m from station			0.056*	0.052*	0.048+	0.043+
			(0.022)	(0.022)	(0.025)	(0.025)
Tract Pair FE	Y	Y	Y	Y	Y	Y
POW-X-Yr FE	Y	Y	Y	Y	Y	Y
RES-X-Yr FE	Y	Y	Y	Y	Y	Y
Sbcty-X-Sbcty-X-Yr FE	-	Y	-	Y	-	Y

High-dimensional fixed effects estimates of λ^D. Treatment variables are mutually exclusive with others in each column. All control networks are 'loose' (see text). All estimates include tract of work-by-year, tract of residence-by-year, and tract pair fixed effects. Standard errors clustered by tract pair, tract of residence, and tract of work in parentheses: $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

For Online Publication

Appendices and Supplemental Results

to accompany

Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification

by

Christopher Severen

October 2019

A Discussion of Data

In this Appendix section, I discuss all the sources of data that this project draws from and details relevant to sample construction. I pay particular attention to normalization. I also compare the CTPP and LEHD LODES data sources, and explain why they are not suitable to be used together.

A.1 Sources

- Census Transportation Planning Project (CTPP)
  - 1990 Urban Part II: Place of Work, Census Tract
  - 1990 Urban Part III: Journey-to-Work, Census Tract
  - 2000 Part 2
  - 2000 Part 3
  - 2006/10 Part 3 (not used in current draft)
- National Historical Geographic Information System (NHGIS)
  - Shapefiles, Block Group and Census Tract, 1990, 2000, and 2010
  - Census, Block Group and Census Tract aggregates, 1990 and 2000
- Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES)
  - Aggregated to tract-to-tract flows, 2002 and 2015, using constant 2010 geographies
- Geolytics Neighborhood Change Database (NCDB)
  - Census aggregates in constant 2010 geographies from 1970-2010
- Los Angeles County Metropolitan Transportation Authority (LACMTA)
  - Shapefiles on LA Metro stations and lines
  - Opening dates for stations and lines
  - Ridership data
  - Kelker, De Leuw & Company (1925). I georeference this map in ArcGIS, and then process it in R to provide geographic data to delineate the 1925 Plan and PER Line samples.
- IPUMS USA
  - Microdata on employment, wage, and industry by MSA for all non-CA residents, 1980-2000.
  - Microdata on transit in the 1990 and 2000 Censuses for LA area residents.
- Southern California Association of Governments
  - Land use and zoning maps: 1990, 1993, 2001, 2005.
- National Highway Planning Network
  - Shapefiles for the Century Freeway (I-105)

A.2 Data construction details

See data map on following pages. Where there have been significant and arbitrary data processing decisions, I denote this by P#.

Geographic normalization

Through my primary analysis (all results from 1990 and 2000, excluding the check on pre-trends), the unit of observation is the census tract according to 1990 Census geographies. The Transportation Analysis Zones used in Southern California in the 1990 CTPP are equivalent to census tracts from the 1990 Census that have been subdivided by municipal boundaries if they overlay multiple jurisdictions. I merge TAZs in 1990 that cross municipal boundaries and assign them to the corresponding census tract. Data from the 2000 CTPP and 2000 Census are both in 2000 geographies. I therefore overlay shapefiles delineating 2000 geographies on 1990 census tracts to develop a crosswalk that translates 2000 data into 1990 geographies.[F.1] Where possible, I use 2000 block group data and shapefiles to refine the crosswalk. More precisely, to create the crosswalk, I intersect the 2000 census tracts and census block group files with 1990 census tracts, and then clean to provide a set of weights to be used in converting 2000 data to the 1990 geographies. Note that the intersection method varies according to whether summation or averaging is desired. If summing, weights are the portion of a 2000 geography that overlays the 1990 census tract. If averaging, weights are the portion of the 1990 census tract that is covered by a 2000 geography. In all cases, I excluded intersected values that cover less than 0.5% of the targeted area to reduce noise (P1).[F.2]

To normalize 2000 flows and travel times to 1990 geographies, the crosswalk is merged twice into the data, once by origin and once by destination (using the Stata command joinby to ensure all combinations were made). I then collapse these data by 1990 origin-destination pairs, taking the raw sum areal weights as the 1990 flow counts and using the areal weights to determine travel times. Many travel times are not disclosed in the 2000 data, and are treated as missing and are ignored. The 2000 CTPP data do not report actual counts, instead rounding to the nearest 5 (except for 1-7, which is labeled 4). In order to treat 1990 and 2000 data similarly, I develop two approaches that are conservative, though they throw away potentially useful variation. Both are similar, but differ in how they treat small numbers. In approach (P2a), I divide flows by 5, and round to the nearest digit. In approach (P2b), I change any flow values between 1 and 4 inclusive to be 4, and divide by 5 and round to the nearest digit. Small digits are different in the two years: In 1990, digits <4 have actual meaning, whereas in 2000 digits <4 can only have been created through the areal weighting process. Both approaches accommodate these differences in a different way, and offer different truncation points (2.5 for approach (a), and 1 for approach (b)). Approach (b) is my preferred specification. For all flows-by-mode, I follow approach (b), as not doing so would result in significant left-truncation. I also drop all pairs with a value of 0 in both 1990 and 2000 for approach (b) (P4). A small number of locations failed to merge. The flows in these amounted to 0.4% of the population.

I exclude census tracts from the eastern edges of San Bernardino and Riverside counties on the Channel Islands.

Labor demand shock construction

I construct wage and employment variants of the Bartik (1991) labor demand shock using Census microdata from 1990 and 2000. I exclude all workers in California. To create measures of national changes in labor demand, I calculate the change in wage or employment by two digit SIC industry from 1990 to 2000. I then interact this with the 1990 employment share by industry at each census tract of work to create a local measure of (plausibly exogenous) change in labor demand. While it would be preferable to use 1980 employment share by industry at tract of work, I have not been able to locate such data.

I then follow the approach described in Section 4 and interact the labor demand shock with the distance between tracts to model how the shock dissipates into adjacent markets. Because each tract may be joined to a different number of tracts, I weight by distance and exclude tracts that experience zero commuting flows (P3).

F.1. This is essentially the reverse process of the Longitudinal Tract Data Base in Logan, Xu, and Stults (2014); I bring current data to 1990 geographies because merging tracts induces less error than (perhaps incorrectly) splitting tracts.

F.2. There are constant small realignments of census blocks (which aggregate to tracts) to account for roads, construction, lot mergers, etc. I choose the 0.5% threshold because it is unlikely that this represented a substantive change in the census tract, but rather just a minor border adjustment.

Data trimming

The various processes above produce relatively standardized data that accord reasonably well with ad hoc probes of quality. However, there are instances of extreme values that become influential observations during estimation. I experimented with a number of approaches to deal with this: (i) doing nothing, (iia) winsorizing in levels, (iib) trimming in levels, (iiia) winsorizing in changes, and (iiib) trimming in changes, where all winsorizing and trimming takes places at the 1st and 99th centiles. I ultimately settled on (iiib) trimming in changes, because it reduces the number of influential observations and removes observations with implausible-seeming characteristics from the data. I also remove observations with top-coded data where applicable. If a variable was top-coded differently in different years, I standardized the top code to the most conservative year.

Construction of treatment and control groups

The Dorothy Peyton Gray Transportation Library of LACMTA hosts historical data on proposed transit plans for the Los Angeles area, including the Kelker, De Leuw & Company (1925) plan. I obtain high-resolution digital copies of Plates 1 and 2 of this document and georeference them in ArcGIS using immutable landmarks and political boundaries.[F.3] I then trace the proposed lines and the existing PER lines from this map, and convert these traces into shapefiles.

To define treatment status, I spatially join shapefiles on actual LA Metro Rail stations from LACMTA to both census tract centroids and boundaries. I define treatment in two ways:

- A narrower definition that requires that either (i) the distance from a tract boundary to a station be exactly 0, or (ii) the distance from a tract centroid to a station be less than 500 meters. Condition (i) implies that the stations lie within the census tract.
- A broader definition that requires just that the distance from a tract boundary to a station be less than 500 meters.

All treated tracts are included in all estimates. To develop a set of control tracts, I spatially join the shapefiles descended from the Kelker, De Leuw & Company (1925) document to the census tract shapefiles, and keep all tracts that have boundaries within 500 meters of the tracks. This assigns non-treated tracts to a control group for three different reasons: (i) They lie along spurs of proposed track that were never built, (ii) they are near a built track but distant from a station, or (iii) they lie slightly farther away from stations than nearby treated tracts. Previous iterations of this paper have used alternative definitions of these control groups, but the use of a 500 meter boundary seems to provide the closest comparison. I perform this separately for 1990 tract geographies (for the main specifications) and 2010 tract geographies (for use with the NCDB and LEHD LODES).

A.3 CTPP vs. LODES

I draw data primarily from the CTPP. There are a number of advantages and a few disadvantages of the CTPP over another popular source of data, the Longitudinal Employer-Household Dynamic (LEHD) Origin-Destination Employment Statistics (LODES). The benefits of CTPP data:

1. In CTPP data, place of work is determined from household responses to a particular set of census questions. The response indicates where an individual worked in the week prior to the census, which may or may not correspond to a fixed establishment. LODES data come from federal tax records, and so identify people as working at the address on a firm's tax statement. Thus for firms with several establishments, there may be clustering at the mailing location that is not indicative of actual workplace. This is particularly true for large, multi-establishment firms.

F.3. Maps available through the LACMTA library and online at https://www.metro.net/about/library/archives/visions-studies/mass-rapid-transit-concept-maps/.

2. The CTPP included median and mean wage at place of work prior in the 1990 and 2000 enumerations. LODES provides only a few large bins. Accurate measures of local wage at place of work are key to this analysis, and a novel contribution to the urban trade literature.
3. CTPP data include reported travel times. Thus, these estimates take into account congestion and other items unobservable to route planning GIS systems that may induce measurement error.
4. CTPP location data are accurately reported, while there is some geographic randomization (within block group) in LODES data to preserve confidentiality.
5. The CTPP data go back to 1990, while LODES does not begin until 2002. Thus, with CTPP I can fully capture commuting in 'pre' and 'post' periods.

Benefits of LODES data:

1. LODES data provide annual measures of commuting between locations since 2002, and the geocoding of workplace mailing address has a higher match rate than in the CTPP.
2. The CTPP has rather odd rounding rules that induce more measurement error in low commute-flow tract pairs. LODES has no such rounding rules (though there is geographic jittering).
3. LODES is calculated with consistent geography over time, while the CTPP is estimated using whatever geographies are decided upon by state census and transportation entities. This means that CTPP data must undergo geographic normalization, while LODES data do not.

There are two further disadvantages to the CTPP data: (i) Not all fields from the 1990 and 2000 CTPP are reported in the 2006/10 CTPP. Important for this paper is the lack of wage at place of work data in 2006/10. (ii) Industry coding changed between the 1990 and 2000 census reports.

I have tried combining data sources to provide a more complete panel of commuting flows across time. There are a number of issues with this approach, namely concern that measurement error in flows drowns out meaningful variation in observed commuting flow changes over time. In fact, this seems to be the case when combining the 1990 CTPP with 2002 LODES data, or the 1990 and 2000 CTPP data with more recent LODES data. Further, the lack of wage at place of work data in LODES is a severe disadvantage. While I have experimented with alternative (fixed effects) methods to estimate wage at place of work, measurement error swamps meaningful measurement.

B Proofs and Algebra

Proposition 1

To establish Proposition 1i (existence), I utilize a fixed point argument and homogeneity. To establish Proposition 1ii, I make use of Theorem 1ii from Allen, Arkolakis, and Li (2014) (AAL) and the Perron-Frobenius Theorem.

Existence in a closed economy: Land use is assumed to be predetermined. Denote the set of location pairs with positive land use for housing and production as $\mathcal{C} = \{ij : L_i^H > 0 \text{ and } L_j^Y > 0\}$, and the cardinality of $\mathcal{C}$ as $N_{\mathcal{C}}$. Assume that $L_i^H > 0 \Leftrightarrow \sum_s \pi_{is} > 0$ and $L_j^Y > 0 \Leftrightarrow \sum_r \pi_{rj} > 0$. The model can be entirely expressed in terms of the aggregate population $\bar{N}$, the data on land use, local fundamentals, travel costs, and commuting shares $\{L_i^H, L_j^Y, A_j, \tilde{B}_i, C_i, D_{ij}, E_j, T_i, \delta_{ij}, \pi_{ij}\}_{\forall ij \in \mathcal{C}}$. Note that the commuting shares and aggregate population are endogenous, all else is given.

The commuting share from ij can be written as an implicit function of the vector of all commuting shares, population, exogenous variables, and models parameters: Define $\mathcal{T}_{ij}(\boldsymbol{\pi}; \bar{N})$:

$$\mathcal{T}_{ij}(\boldsymbol{\pi}; \bar{N}) = \frac{\frac{\Lambda_{ij}}{\delta_{ij}^{\epsilon}} \cdot \frac{\breve{A}_j^{\epsilon}}{\left(\bar{N}\sum_r \pi_{rj}\right)^{\epsilon(1-\alpha)}} \cdot \left(\bar{N}\breve{C}_i \cdot \sum_s \frac{\pi_{is}\breve{A}_s}{\left(\bar{N}\sum_r \pi_{rs}\right)^{1-\alpha}}\right)^{\frac{-\epsilon\psi(1-\zeta)}{1+\psi}}}{\sum_r \sum_s \frac{\Lambda_{rs}}{\delta_{rs}^{\epsilon}} \cdot \frac{\breve{A}_s^{\epsilon}}{\left(\bar{N}\sum_{r'} \pi_{r's}\right)^{\epsilon(1-\alpha)}} \cdot \left(\bar{N}\breve{C}_r \cdot \sum_{s'} \frac{\pi_{rs'}\breve{A}_{s'}}{\left(\bar{N}\sum_{r'} \pi_{r's'}\right)^{1-\alpha}}\right)^{\frac{-\epsilon\psi(1-\zeta)}{1+\psi}}}$$

with $\breve{A}_j = \alpha A_j L_j^{Y\,1-\alpha}$ and $\breve{C}_i = (1-\zeta) C_i^{1/\psi} L_i^{H\,-1}$. An equilibrium of the model is the vector $\boldsymbol{\pi}$ and aggregate population $\bar{N}$ such that $\boldsymbol{\pi}$ is a fixed point of $\mathcal{T}_{ij}(\boldsymbol{\pi}; \bar{N})$ and the no spatial arbitrage condition is satisfied. First, note that $\mathcal{T}_{ij}(\boldsymbol{\pi}; \bar{N})$ is homogeneous of degree zero in $\bar{N}$, so $\mathcal{T}_{ij}(\boldsymbol{\pi}; \bar{N}) = \mathcal{T}_{ij}(\boldsymbol{\pi})$ and the existence of commuting shares is independent of aggregate population.

Consider $\mathcal{T}_{ij}(\boldsymbol{\pi})$. By assumption, for all $ij \in \mathcal{C}$, we have $L_i^H > 0$, $L_j^Y > 0$, and $\sum_r \pi_{rj} > 0$ and $\sum_s \pi_{is} > 0$. This implies that $\pi_{ij} \geq 0$, and $\pi_{ij} \leq 1$ because π represent shares. Stacking equations, equilibrium commuting shares are a fixed point $\mathcal{T}(\boldsymbol{\pi}^{FP}) = \boldsymbol{\pi}^{FP}$. The function $\mathcal{T} : [0,1]^{N_{\mathcal{C}}} \to [0,1]^{N_{\mathcal{C}}}$ is continuous and maps a compact, convex set into itself. Therefore, by the Brouwer fixed point theorem, an equilibrium vector $\boldsymbol{\pi}^{FP}$ exists. In a closed economy, aggregate population is fixed, so this establishes existence.

Existence in an open economy: In an open economy, existence of equilibrium follows from *Existence in a closed economy*, but also the no spatial arbitrage that requires expected utility to be equalized to $\bar{U}$ in equilibrium. Denote element ij of $\boldsymbol{\pi}^{FP}$ be π_{ij}. Rewriting the no spatial arbitrage condition:

$$\bar{N} = \left(\frac{\bar{U}}{\Gamma\left(\frac{\epsilon-1}{\epsilon}\right) \cdot \left(\sum_r \sum_s \frac{\Lambda_{rs}}{\delta_{rs}^{\epsilon}} \cdot \frac{\breve{A}_s^{\epsilon}}{\left(\bar{N}\sum_{r'} \pi_{r's}\right)^{\epsilon(1-\alpha)}} \cdot \left(\bar{N}\breve{C}_r \cdot \sum_{s'} \frac{\pi_{rs'}\breve{A}_{s'}}{\left(\bar{N}\sum_{r'} \pi_{r's'}\right)^{1-\alpha}}\right)^{\frac{-\epsilon\psi(1-\zeta)}{1+\psi}}\right)^{1/\epsilon}}\right)^{\frac{1}{1-\alpha\left(1+\frac{\psi(1-\zeta)}{1+\psi}\right)}}$$

Given $\boldsymbol{\pi}^{FP}$, existence requires that the preceding equation give a real, finite value of $\bar{N}$. This is the case so long as $\epsilon > 1$ and $\alpha \neq \frac{1+\psi}{1+\psi(2-\zeta)}$.

Uniqueness: Consider now the set of places with the either positive land use for housing or for production, denoted $\mathcal{J}$ (a theorem referenced below requires that the set of possible housing locations be the same as the set of possible production locations). Rearranging the system in Equations (1), (8), (10), (11), and (12) into a more convenient form gives:

$$W_j^{\frac{1+\epsilon(1-\alpha)}{1-\alpha}}\Omega_j = \bar{N}^{-1}K_{0j}\sum_{s\in\mathcal{J}} W_s^{\epsilon}\Omega_s$$

$$\Omega_j = \sum_{r\in\mathcal{J}} K_{1rj}Q_r^{-\epsilon(1-\zeta)}$$

$$Q_i^{-\epsilon(1-\zeta)-\frac{1+\psi}{\psi}}\Phi_i = \bar{N}^{-1}K_{2i}\sum_{s\in\mathcal{J}} W_s^{\epsilon}\Omega_s$$

$$\Phi_i = \sum_{r\in\mathcal{J}} K_{1is}W_s^{\epsilon+1}$$

where $K_{0j} = \breve{A}_j^{1/(1-\alpha)}$, $K_{1ij} = \Lambda_{ij}\delta_{ij}^{-\epsilon}$, and $K_{2i} = \breve{C}_i^{-1/\psi^2}$ are functions of predetermined parameters.

This transforms the model into the form of Equation 1 in AAL. Let $\mathbb{G}$ represent the matrix of exponents on the left hand side of the above system in the order (W, Ω, Q, Φ), and let $\mathbb{B}$ be the corresponding exponents on the right hand side:

$$\mathbb{G} = \begin{pmatrix} \frac{1+\epsilon(1-\alpha)}{1-\alpha} & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\epsilon(1-\zeta)-\frac{1+\psi}{\psi} & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} \epsilon & 1 & 0 & 0 \\ 0 & 0 & -\epsilon(1-\zeta) & 0 \\ \epsilon & 1 & 0 & 0 \\ \epsilon+1 & 0 & 0 & 0 \end{pmatrix}$$

Note that $\mathbb{G}$ is invertible. To address uniqueness, define $\mathbb{A} = \mathbb{B}\mathbb{G}^{-1}$ and $\mathbb{A}^+$ to be the element-wise absolute value of $\mathbb{A}$. That is,

$$\mathbb{A}^+ = \begin{pmatrix} \frac{\epsilon[(1-\alpha)-\mu]}{1+\epsilon(1-\alpha)} & \frac{1}{1+\epsilon(1-\alpha)} & 0 & 0 \\ 0 & 0 & \frac{\epsilon(1-\zeta)}{\epsilon(1-\zeta)+\frac{1+\psi}{\psi}} & \frac{\epsilon(1-\zeta)}{\epsilon(1-\zeta)+\frac{1+\psi}{\psi}} \\ \frac{\epsilon[(1-\alpha)-\mu]}{1+\epsilon(1-\alpha)} & \frac{1}{1+\epsilon(1-\alpha)} & 0 & 0 \\ \frac{(\epsilon+1)[(1-\alpha)-\mu]}{1+\epsilon(1-\alpha)} & \frac{(\epsilon+1)(1-\alpha)}{1+\epsilon(1-\alpha)} & 0 & 0 \end{pmatrix}$$

Theorem 1ii in AAL establishes that there is a unique equilibrium to the model if the spectral radius (largest eigenvalue) of $\mathbb{A}^+$ is less than or equal to one. Thus, uniqueness is established when $\rho(\mathbb{A}^+) \leq 1$.

Because $\mathbb{A}^+$ corresponds to a strongly connected graph and is nonnegative, it is irreducible. The Perron-Frobenius Theorem states that a nonnegative, irreducible matrix has a positive spectral radius with corresponding strictly positive eigenvector. So finding a condition under which $\rho(\mathbb{A}^+) \leq 1$ is identical to determining conditions under which $\mathbb{A}^+\mathbf{x} \leq \mathbf{x}$ for $\mathbf{x} \gg 0$. Solving the implied system of inequalities gives condition (13).[B.1]

Proposition 2

<u>Existence:</u> A_i is uniquely determined from:[B.2]

$$A_i = \frac{W_i}{\alpha}\left(\frac{\sum_r \bar{N}\pi_{ri}}{L_i^Y}\right)^{1-\alpha}$$

B.1. To ensure the algebra is correct, I have numerically verified $\rho(\mathbb{A}^+) \leq 1$ iff Equation (13) holds.

B.2. Uniqueness of A holds under agglomeration, the other terms are unaffected.

and C_i is uniquely determined from:

$$C_i = Q_i^{1+\psi}\left(\frac{L_i^H}{\sum_s \bar{N}\pi_{is}W_s}\right)^{\psi}$$

Define an excess demand function:

$$\mathcal{D}_{ij}(\mathbf{\Lambda}) = \pi_{ij} - \frac{\Lambda_{ij}W_j^{\epsilon}\left(\delta_{ij}Q_i^{1-\zeta}\right)^{-\epsilon}}{\sum_r\sum_s \Lambda_{rs}W_s^{\epsilon}\left(\delta_{rs}Q_r^{1-\zeta}\right)^{-\epsilon}} = 0$$

Note that $\mathcal{D}$ is continuous and homogeneous of degree zero. Homogeneity implies that $\mathbf{\Lambda}$ can be rescaled and restricted to the unit simplex: $\{\mathbf{\Lambda} : \sum_r\sum_s \Lambda_{rs} = 1\}$. This means that $\mathcal{D} : [0,1]^{N^2} \to [0,1]^{N^2}$. So $\mathcal{D}$ is a continuous function from a compact, convex set into itself; the Brouwer fixed point theorem guarantees existence.

Uniqueness: To establish uniqueness, note that by homogeneity of degree zero, we have $\sum_r\sum_s \mathcal{D}_{rs}(\mathbf{\Lambda}) = 0$. Define $M_{ij} = W_j^{\epsilon}\left(\delta_{ij}Q_i^{1-\zeta}\right)^{-\epsilon}$. The Jacobian of $\mathcal{D}$ has diagonal elements:

$$-\frac{M_{ij}\cdot\left(\left(\sum_r\sum_s \Lambda_{rs}M_{rs}\right) - \Lambda_{ij}M_{ij}\right)}{\left(\sum_r\sum_s \Lambda_{rs}M_{rs}\right)^2} < 0$$

and off-diagonal elements

$$\frac{\Lambda_{ij}M_{ij}M_{\{ij\}'}}{\left(\sum_r\sum_s \Lambda_{rs}M_{rs}\right)^2} > 0$$

where $\{ij\}'$ refers to an origin destination pair such that $i' \neq i$ and/or $j' \neq j$. Thus the aggregate excess demand function exhibits gross substitution, and equilibrium is unique.[B.3]

Rewriting the Model as a Three Linear Equation System

Taking logs of Equations (1), (8), and (10) gives the following cross-sectional system (where lowercase letters represent the log counterparts of level variables):

$$w_j = g_0 + (\alpha - 1)n_j^Y + \ln(A_j) \tag{1}$$

$$n_{ij} = g_1 + \epsilon w_j - \epsilon(1-\zeta)q_i - \epsilon\kappa\tau_{ij} + \ln(B_iE_jD_{ij}) \tag{2}$$

$$q_i = g_2 + \psi h_i + \ln(C_i) \tag{3}$$

where $n_j^Y = \ln\left(\bar{N}\sum_r \pi_{rj}/L_j^Y\right)$ is log employment density and $h_i = \ln\left((1-\zeta)\bar{N}\sum_s \pi_{is}W_s/Q_iL_i^H\right)$ is log housing density. The g capture remaining constants: $g_0 = \ln(\alpha)$, $g_1 = \ln(\bar{N}) - \ln\left(\sum_r\sum_s \Lambda_{rs}\left(e^{\kappa\tau_{rs}}Q_r^{1-\zeta}\right)^{-\epsilon}W_s^{\epsilon}\right)$, and $g_2 = 0$. Local fundamentals are potentially functions of covariates ($A = A(X)$ and so on) such as transit proximity.

This system can be re-expressed to more clearly represent the supply and demand linkages and better exposit the identification strategy. First, separate the unobservables into time-varying and time-invariant components, so that $\ln(A_{jt}) = \bar{a}_j + a_{jt}$, etc. Under the assumption that land use and travel times are

B.3. See Proposition 17.F.3 in Mas-Colell, Whinston, and Green, *Microeconomic Theory* (Oxford University Press, 1995). An alternative approach could be to use weak diagonal dominance of this positive matrix (following Bayer and Timmins (2005) but for weaker conditions).

constant, this means making the structural assumptions:

$$\ln(A_{jt} L_{jt}^{Y\,1-\alpha}) = \bar{a}_j + a_{jt} \tag{4}$$
$$\ln(B_{it} E_{jt} D_{ijt}) = \bar{b}_i + b_{it} + \bar{e}_j + e_{jt} + \bar{d}_{ij} + d_{ijt} \tag{5}$$
$$\ln(C_{it} L_{it}^{H-\psi}) = \bar{c}_i + c_{it} \tag{6}$$

Relaxing this to allow for exogenous changes in land use is straightforward. Doing so, and preserving the notation above leads to the following system:

$$\begin{aligned}
\text{Labor demand in } j\text{:} \qquad & w_{jt} = g_{0t} + \tilde{\alpha} n_{jt}^{Y} + \bar{a}_j + a_{jt} && (7)\\
\text{Labor supply to } j\text{:} \qquad & \omega_{jt} = \epsilon w_{jt} + \bar{e}_j + e_{jt} && (8)\\
\text{Commuting between } i \text{ and } j\text{:} \qquad & n_{ijt} = g_{1t} + \omega_{jt} + \theta_{it} - \epsilon\kappa\tau_{ijt} + \bar{d}_{ij} + d_{ijt} && (9)\\
\text{Housing demand in } i \qquad & \theta_{it} = \tilde{\zeta} q_{it} + \bar{b}_i + b_{it} && (10)\\
\text{Housing supply in } i\text{:} \qquad & q_{it} = g_{2t} + \psi h_{it} + \bar{c}_i + c_{et} && (11)
\end{aligned}$$

where $\tilde{\alpha} = \alpha - 1$, $\tilde{\zeta} = -\epsilon(1-\zeta)$. The system resembles standard linear supply and demand models, but for many interconnected housing and labor markets.

Welfare under $\epsilon \leq 1$ (Frechet is Multinomial Logit)

First, I show that the expression in Equation (18) has an equivalent log-sum representation. Begin by dividing counterfactual and factual expected utilities (from Equation 6):

$$\hat{U} = \frac{\mathbb{E}[U'_{ijo}]}{\mathbb{E}[U_{rso}]} = \frac{\Gamma\left(\frac{\epsilon-1}{\epsilon}\right) \cdot \left(\sum_{\{ij\}} \tilde{\Lambda}'_{ij} \left(\delta'_{ij} Q_i'^{\,1-\zeta}\right)^{-\epsilon} (\tilde{B}'_i W'_j)^{\epsilon}\right)^{1/\epsilon}}{\Gamma\left(\frac{\epsilon-1}{\epsilon}\right) \cdot \left(\sum_{\{rs\}} \tilde{\Lambda}_{rs} \left(\delta_{rs} Q_r^{1-\zeta}\right)^{-\epsilon} (\tilde{B}_r W_s)^{\epsilon}\right)^{1/\epsilon}} = \left(\frac{\sum_{\{ij\}} \tilde{\Lambda}'_{ij} \left(\delta'_{ij} Q_i'^{\,1-\zeta}\right)^{-\epsilon} (\tilde{B}'_i W'_j)^{\epsilon}}{\sum_{\{rs\}} \tilde{\Lambda}_{rs} \left(\delta_{rs} Q_r^{1-\zeta}\right)^{-\epsilon} (\tilde{B}_r W_s)^{\epsilon}}\right)^{1/\epsilon} \tag{12}$$

where $\{ij\} = \{rs\}$ track summation sets. Substituting in Equation (1) for some particular ij into the above twice (once for π'_{ij} and once for π_{ij}) and taking logs gives Equation (18).

From Train (2009), the change in consumer welfare due to changes of the characteristics of the elements in the choice set is:

$$\mathbb{E}[\bar{\mathcal{W}}'] - \mathbb{E}[\bar{\mathcal{W}}] = \frac{1}{\mu} \ln\left(\frac{\sum_{k\in K_1} e^{V'_k}}{\sum_{k\in K_0} e^{V_k}}\right) \tag{13}$$

where here μ is the marginal utility of income.[B.4] Let:

$$
\begin{aligned}
V'_k &= \ln\left(\tilde{\Lambda}'_{ij}\left(\delta'_{ij}{Q'_i}^{1-\zeta}\right)^{-\epsilon}(\tilde{B}'_i W'_j)^{\epsilon}\right)\\
V_k &= \ln\left(\tilde{\Lambda}_{rs}\left(\delta_{rs}{Q_r}^{1-\zeta}\right)^{-\epsilon}(\tilde{B}_r W_s)^{\epsilon}\right)\\
\mu &= \epsilon\\
K_0 = K_1 &= \{ij\} = \{rs\}
\end{aligned}
$$

Taking logs of Equation F-1 then delivers Equation F-2. Note that $\mu = \epsilon$ is natural as ϵ already captures the utility effect of wage dollars. Thus the Frechet framework is identical to a multinomial logit framework where the utility from choice ij is:

$$
\mathcal{U}_{ijo} = \ln\left(\tilde{\Lambda}'_{ij}\left(\delta'_{ij}{Q'_i}^{1-\zeta}\right)^{-\epsilon}(\tilde{B}'_i W'_j)^{\epsilon}\right) + \varepsilon_{ijo}
$$

for ϵ_{ijo} distributed iid extreme value. In fact, this is very precisely (up to interpretation of amenity terms and trade costs) the specification often used in the discrete location choice literature (e.g., Bayer, Keohane, and Timmins 2009). To map interpretation of the change in consumer welfare between the two frameworks, note:

$$
\mathbb{E}[\bar{\mathcal{W}}'] - \mathbb{E}[\bar{\mathcal{W}}] = \ln \mathbb{E}[U'_{ijo}] - \ln \mathbb{E}[U_{ijo}] = \ln \hat{\bar{U}} \approx \%\Delta \text{ Welfare}
$$

That is, welfare change is naturally expressed in relative terms (rather than monetary terms) when used with Frechet framework. Equation F-2 only requires $\epsilon > 0$, and so Equation (18) can be used for welfare evaluation when $\epsilon \in (0, 1]$ and well as $\epsilon > 1$.

B.4. Thanks to Wei You for noting that (18) and a log-sum expression are interchangeable:

$$
\hat{\bar{U}} = \left(\frac{\hat{\bar{\Lambda}}_{ij}(\hat{\bar{B}}_i \hat{W}_j)^{\epsilon}\hat{Q}_i^{-\epsilon(1-\zeta)}}{\hat{\pi}_{ij}}\right)^{1/\epsilon} = \left(\sum_{\{ij\}} \pi_{ij}\hat{\bar{\Lambda}}_{ij}\left(\hat{\delta}_{ij}\hat{Q}_i^{1-\zeta}\right)^{-\epsilon}(\hat{\bar{B}}_i \hat{W}_j)^{\epsilon}\right)^{1/\epsilon}
$$

C Cost-benefit Calculations

This section details the costs of the subway built by 2000. I do not track costs since 2000, as the calculation becomes much less clear with more recent data. To compare the costs and benefits of transportation interventions, I require annualized estimates of costs to compare with the annualized welfare benefits calculated in the text. Costs consist of two components: (i) the annualized cost of capital investment in rail, rail cars, stations, and similar expenses, and (ii) net operating expenses (operating costs less revenues).

$$\text{Total Annual Cost} = \text{Operating Subsidy} + \text{Annualized Capital Expenditure}$$

C.1 Annualized Capital Expenditure

Cost information is from a consolidation of capital expenditures on lines built before 2000 from fiscal budgets.[C.1] After adjusting all costs to 2015 dollars, the total capital expenditure for the rail, rolling stock, and stations built prior to 2000 is $8.7 billion. To annualized this, I assume annual payments are made on this principal balance over a 30-year horizon with 6% interest rate (the interest rate used for some internal calculations by LA Metro). This gives an annualized capital cost of $634.6 million. This does not include other financing charges, the cost of planning, or some other expenses.

However, LA Metro's internal cost of borrowing may not be a suitable social discount rate, and the 30-year horizon may be too short. I provide several alternative definitions: (i) 5% interest over a 50-year horizon, (ii) 5% over an infinite horizon, and (iii) 2.5% over an infinite horizon. For (i) and (ii), the 5% rate is roughly equal to a low-yielding municipal bonds in 2000. For (iii), the 2.5% rate is low, roughly equal to the recent cost of borrowing, and is meant to represent a policy maker that highly values future generation or is uncertain about future discount rates (see Weitzman 1998). Once built, subways typically remain in operation for the long run (perhaps forever).

C.2 Operating Subsidies

Like most transit systems in the United States, LA Metro has incomplete farebox recovery, meaning that it subsidizes a portion of every ride. For rail in 2001, the farebox recovery ratio was about 20%. To estimate the welfare effects, I use the *net* subsidy: operating costs less fare revenue. Operating expenses from 1999 or 2000 are unavailable, so I use operating expenses from 2001 and 2002 as a proxy. Rail (light and heavy) operations total $202.4 million in 2015 dollars, and rail fare revenue is $40.2 million. The net subsidy is $162.2 million per year.

C.3 Sales Taxes

Most of the funding for the early stages of rail construction in Los Angeles came from federal transportation grants, and so I ignore any deadweight loss imposed by local sales tax in the primary analysis. However, as robustness, I perform some deadweight loss calculations. Because preferences are Cobb-Douglas, a constant share of income is spent on consumption, and so adding sales tax on the composite consumption goods does not have any direct effect on household housing expenditures.

To incorporate the tax, I apply the following schedule to Los Angeles County residents, making the simplifying assumption that households purchase the consumption good where they live. The primary taxes that funded early stages of rail development were Proposition A (1980) and Proposition C (1990). Proposition A was a half-cent increase with 35% earmarked for subway and light rail, while Proposition C was a half-cent increase with 45% allocated to subway and light rail. The base right prior is assumed at 6% (the rate in Los Angeles prior to 1980). Therefore, I model transit sales tax as an increase from 6% to 6.4% for Los Angeles County residents.

C.1. Source: http://demographia.com/db-rubin-la-transit.pdf.

Implementing this in the model is straightforward: Indirect utility with Sales Tax $\bar{ST}_i$ in location i is:

$$v_{o|ij} = \frac{\nu_{ijo}\tilde{B}_i W_j Q_i^{\zeta-1}\bar{ST}_i^{-\zeta}}{\delta_{ij}}$$

Population shares that choose residential location i and place of work j are then:

$$\pi_{ij} = \frac{\tilde{\Lambda}_{ij}\left(\delta_{ij}Q_i^{1-\zeta}\bar{ST}_i^{\zeta}\right)^{-\epsilon}(\tilde{B}_i W_j)^{\epsilon}}{\sum_r \sum_s \tilde{\Lambda}_{rs}\left(\delta_{rs}Q_r^{1-\zeta}\bar{ST}_r^{\zeta}\right)^{-\epsilon}(\tilde{B}_r W_s)^{\epsilon}} \tag{1}$$

See the next section for counterfactual estimation. Note that, were the tax to apply uniformly across the entire region, the effect could be estimated simply and analytically.

D Counterfactual Estimation

First, note that the following hold:

$$
\begin{aligned}
\hat{W}_i &= \hat{A}_i \hat{\hat{N}}^{\alpha-1} \left(\frac{\sum_r \pi_{ri} \hat{\pi}_{ri}}{\sum_r \pi_{ri}} \right)^{\alpha-1} \\
\hat{Q}_i &= \hat{C}_i^{1/(1+\psi)} \left(\frac{\hat{\hat{N}} \sum_s \pi_{is} \hat{\pi}_{is} W_s \hat{W}_s}{\sum_s \pi_{is} W_s} \right)^{\psi/(1+\psi)} \\
\hat{\pi}_{ij} &= \frac{\hat{B}_i \hat{E}_j \hat{D}_{ij} \hat{W}_j^{\epsilon} \hat{Q}_i^{-\epsilon(1-\zeta)}}{\sum_r \sum_s \pi_{rs} \hat{B}_r \hat{E}_s \hat{D}_{rs} \hat{W}_s^{\epsilon} \hat{Q}_r^{-\epsilon(1-\zeta)}}
\end{aligned}
$$

where $\hat{\hat{N}} = 1$ in a closed economy. In the case of the open economy, aggregate population can adjust, ensuring no arbitrage between the city and outside locations. To account for this, define:

$$
\hat{\hat{N}} = \left(\sum_r \sum_s \pi_{rs} \hat{B}_r \hat{D}_{rs} \left(\hat{A}_s \left(\frac{\sum_{r'} \pi_{r's} \hat{\pi}_{r's}}{\sum_{r'} \pi_{r's}} \right)^{\alpha-1} \right)^{\epsilon} \times \left(\hat{C}_r \cdot \left(\frac{\sum_{s'} \pi_{rs'} \hat{\pi}_{rs'} W_{s'} \hat{A}_{s'} \left(\frac{\sum_{r'} \pi_{r's} \hat{\pi}_{r's}}{\sum_{r'} \pi_{r's}} \right)^{\alpha-1}}{\sum_{s'} \pi_{rs'} W_{s'}} \right)^{\psi} \right)^{\frac{-\epsilon(1-\zeta)}{1+\psi}} \right)^{\frac{1+\psi}{\epsilon[(1+\psi)-\alpha(1+\zeta\psi)]}}
$$

Simulating counterfactuals

Simulate closed economy counterfactual, then use that as the initial guess for the open economy counterfactual:

1. Make an initial guess of wages and housing prices: $\{\hat{W}_i^{(0)}\}$, $\{\hat{Q}_i^{(0)}\}$. It is useful to set these equal to 1.
2. Estimate $\{\hat{\pi}_{ij}^{(0)}\}$ using $\{\hat{W}_i^{(0)}\}$, $\{\hat{Q}_i^{(0)}\}$, and $\{\pi_{ij}\}$.
3. Main Loop:
   (a) Define $\{\hat{Q}_i^{(temp)}\}$ using $\{\hat{W}_i^{(t-1)}\}$, $\{W_i\}$, $\{\hat{\pi}_{ij}^{(t-1)}\}$, and $\{\pi_{ij}\}$.
   (b) Define $\{\hat{W}_i^{(temp)}\}$ using $\{\hat{\pi}_{ij}^{(t-1)}\}$, and $\{\pi_{ij}\}$.
   (c) Define $\{\hat{\pi}_{ij}^{(temp)}\}$ using $\{\hat{W}_i^{(t)}\}$, $\{\hat{Q}_i^{(t)}\}$, and $\{\pi_{ij}\}$.
   (d) Update $\hat{X}^{(t)} = \xi \hat{X}^{(temp)} + (1-\xi)\hat{X}^{(t-1)}$ for $\hat{X} \in \{\hat{Q}, \hat{W}, \hat{\pi}$, where ξ is a weight that disciplines updating.
   (e) Estimate movement as:
   $$\Delta = \sum_r |\hat{W}_r^{(t)} - \hat{W}_r^{(t-1)}| + \sum_r |\hat{Q}_r^{(t)} - \hat{Q}_r^{(t-1)}| + \frac{1}{N} \sum_r \sum_s |\hat{\pi}_{rs}^{(t)} - \hat{\pi}_{rs}^{(t-1)}|.$$
   (f) Stop when movement is below convergence criterion.
4. Initial guess for $\hat{\hat{N}}^{(0)}$ using $\{\hat{W}_i^{(temp)}\}$, $\{W_i\}$, $\{\hat{Q}_i^{(temp)}\}$, $\{\hat{\pi}_{ij}^{(temp)}\}$, and $\{\pi_{ij}\}$

5. Main Loop:

   (a) Define $\{\hat{Q}_i^{(temp)}\}$ using $\hat{\bar{N}}^{(t-1)}$, $\{\hat{W}_i^{(t-1)}\}$, $\{W_i\}$, $\{\hat{\pi}_{ij}^{(t-1)}\}$, and $\{\pi_{ij}\}$.

   (b) Define $\{\hat{W}_i^{(temp)}\}$ using $\hat{\bar{N}}^{(t-1)}$, $\{\hat{\pi}_{ij}^{(t-1)}\}$, and $\{\pi_{ij}\}$.

   (c) Define $\{\hat{\pi}_{ij}^{(temp)}\}$ using $\{\hat{W}_i^{(t)}\}$, $\{\hat{Q}_i^{(t)}\}$, and $\{\pi_{ij}\}$.

   (d) Define $\hat{\bar{N}}^{(temp)}$ using $\{\hat{W}_i^{(t)}\}$, $\{W_i\}$, $\{\hat{Q}_i^{(t)}\}$, $\{\hat{\pi}_{ij}^{(t)}\}$, and $\{\pi_{ij}\}$.

   (e) Update $\hat{X}^{(t)} = \xi \hat{X}^{(temp)} + (1-\xi)\hat{X}^{(t-1)}$ for $\hat{X} \in \{\hat{Q}, \hat{W}, \hat{\pi}, \hat{\bar{N}}\}$, where ξ is a weight that disciplines updating.

   (f) Estimate movement as:

$$\Delta = \sum_r |\hat{W}_r^{(t)} - \hat{W}_r^{(t-1)}| + \sum_r |\hat{Q}_r^{(t)} - \hat{Q}_r^{(t-1)}| + \frac{1}{N} \sum_r \sum_s |\hat{\pi}_{rs}^{(t)} - \hat{\pi}_{rs}^{(t-1)}| + |\hat{\bar{N}}^{(t)} - \hat{\bar{N}}^{(t-1)}|.$$

   (g) Stop when movement is below convergence criterion.

With sales taxes

Everything is very similar with local sales taxes, except the ratio of population shares is:

$$\hat{\pi}_{ij} = \frac{\hat{B}_i \hat{E}_j \hat{D}_{ij} \hat{W}_j^{\epsilon} \hat{Q}_i^{-\epsilon(1-\zeta)} \ \hat{\bar{S}T}_i^{-\epsilon\zeta}}{\sum_r \sum_s \pi_{rs} \hat{B}_r \hat{E}_s \hat{D}_{rs} \hat{W}_s^{\epsilon} \hat{Q}_r^{-\epsilon(1-\zeta)} \ \hat{\bar{S}T}_r^{-\epsilon\zeta}}$$

$\hat{\bar{S}T}_i = 1.064/1.06$ for i in Los Angeles County.

E Identification under Additional and Alternative Assumptions

This section extends the approach of interacting labor demand shocks with geography to identify the remaining (housing and labor demand) elasticities. I also discuss two modifications to the standard identification framework: (i) endogenous land use determination (no zoning), and (ii) the presence of agglomeration and other forces.

Residents of one location commute to many different locations for work. Workers who live in i and work in j are sensitive to the housing demands of workers who work in j' but also live in i. A labor demand shock to workers ij' can change the effective housing supply to workers ij. Thus labor demand shocks for ij' workers can be used to instrument changes in housing prices for ij workers and identify the slope of housing demand. To develop an average measure of the shocks for $ij', j' \neq j$, I employ inverse weighting as before, but exclude own tract j:

$$\Delta z^{HS}_{i(-j)t}(\rho) = \sum_{s \neq j} \frac{e^{-\rho\delta_{is}} 1_{\check{n}_{is}>0} \Delta z_{st}}{\sum_{s\neq j} e^{-\rho\delta_{is}} 1_{\check{n}_{is}>0}}$$

Note that place of work-by-year fixed effects (ω_{jt}) control for changes in workplace amenities. The following moment condition identifies $\tilde{\zeta} = \epsilon(1-\zeta)$:

$$\mathbb{E}[\Delta z^{HS}_{i(-j)t}(\rho) \times (\Delta b_{it} + \Delta d_{ijt})] = 0, \; \forall \, i, j' \neq j \tag{M-3}$$

This instrument varies for every commuting pair. It is generally difficult to recover estimates of housing demand without microdata due to difficulties in quantifying housing services. Nonetheless, because tract pairs express more variation than individual tracts, this approach can identify the housing demand elasticity.

Finally, workers employed at j observe the labor demand shock to $j' \neq j$, and may respond by leaving j for j'. This suggests that a labor demand shock at j' can be used to instrument changes in employment at j, functioning as a labor supply shock in j and identifying labor demand. But this is reflected through residential location, rather than through location at place of work. Consider residents in i: A positive shock to j' entices more workers from i the closer j' is to i, rather than the closer j' is to j. The following weighting uses this intuition and interacts with distance twice:

$$\Delta z^{LS}_{jt}(\rho) = \sum_{r} \left(\frac{e^{-\rho\delta_{rj}} 1_{\check{n}_{rj}>0}}{\sum_r e^{-\rho\delta_{rj}} 1_{\check{n}_{rj}>0}} \sum_{s\neq j} \frac{e^{-\rho\delta_{sr}} 1_{\check{n}_{is}>0} \Delta z_{st}}{\sum_{s\neq j} e^{-\rho\delta_{sr}} 1_{\check{n}_{is}>0}} \right)$$

The own tract labor demand shock is excluded in order to remove mechanical correlation with local changes in productivity. The corresponding moment condition is:

$$\mathbb{E}[\Delta z^{LS}_{jt}(\rho) \times \Delta a_{jt})] = 0, \; \forall j \tag{M-4}$$

This identifies the labor demand elasticity, $\tilde{\alpha} = \alpha - 1$, and provides an alternative way to estimate this parameter that is conceptually similar to the competing characteristics instrument of Berry, Levinsohn, and Pakes (1995).

Because the instruments described above are all weighted averages of the labor demand shock, the identifying assumptions can be made more transparent. The following reframe M-1 through M-4 in terms of a labor demand shock (note A-1 is identical to M-1):

$$\mathbb{E}[\Delta z_{jt} \times (\Delta e_{jt} + \Delta d_{ijt})] = 0,\ \forall\, ij \tag{A-1}$$
$$\mathbb{E}[\Delta z_{jt} \times \Delta c_{it}] = 0,\ \forall\, ij \tag{A-2}$$
$$\mathbb{E}[\Delta z_{j't} \times (\Delta b_{it} + \Delta d_{ijt})] = 0,\ \forall\, ij' \neq ij \tag{A-3}$$
$$\mathbb{E}[\Delta z_{j't} \times \Delta a_{jt}] = 0,\ \forall\, j' \neq j \tag{A-4}$$

Proposition 3. *Assume A1, A2, A3, and A4 are true, $\rho > 0$, $\mathbb{E}[\Delta z_{jt} \times \Delta w_{jt}] \neq 0$, housing demand is downward sloping, and labor and housing supply are upward sloping. Then M1, M2, M3, and M4 are satisfied and the model is identified.*

Proof. Assumptions A-1 to A-4 are derived from M-1 to M-4 using the definitions of the instruments. The requirement that $\rho > 0$ ensures variation in the labor demand shock across space. The requirements are standard regularity conditions for identification in a system of simultaneous equations. □

Furthermore, data on commuting flows and workplaces wages in combination with Equation (15) suggest high-dimensional fixed effects can help control for unobserved confounders. Assumptions A-1 and A-3 can be weakened to exploit this:

$$\mathbb{E}[\Delta z_{jt} \times \Delta e_{jt}] = 0,\ \forall\, j \tag{A-1a}$$
$$\mathbb{E}[\Delta z_{jt} \times \Delta c_{it}] = 0,\ \forall\, i \neq j \tag{A-2a}$$
$$\mathbb{E}[\Delta z_{j't} \times \Delta b_{it}] = 0,\ \forall\, i \tag{A-3a}$$

It is difficult to estimate household expenditure shares or labor demand elasticities in urban models that use aggregated data (e.g., Diamond 2016). However, estimates in Tables E1 and E2 roughly concur with values derived from other sources, lending additional credibility to the identification strategy as a whole. Table E1 uses the employment variant of $\Delta z^{HS}_{i(-j)t}$ to instrument for housing prices to determine $\epsilon(1-\zeta)$, the elasticity of housing demand. The own tract can be excluded from the regression to limit concerns about the labor demand shock driving confounding changes in amenities. Results are marginally significant and vary between -0.66 and -0.87. With $\epsilon = 1.83$, these imply a housing expenditure share between 36% and 48% of income, somewhat higher than microdata suggest but not unreasonable for high cost areas.

Table E1: IV estimates of housing demand elasticity ($-\epsilon(1-\zeta)$)

	$n_{ijt} = \ln(N_{ijt})$		
	(1)	(2)	(3)
$\ln$(House Value)	-0.662^{+}	-0.659^{+}	-0.871*
	(0.353)	(0.353)	(0.356)
F-stat (KP)	260.99	261.28	258.84
Sample	All	All	not ii
Travel Time	-	Y	-
N	287598	287598	282754

Panel instrument variable (IV) estimates of regression of flows on median housing values, using $\ln(\rho) = -5.5$. Estimated in differences using employment instrument. KP refers to the Kleinbergen-Papp F-statistic. Variables are trimmed to exclude extreme values (see text). All estimates include tract-of-work-by-year and tract-pair fixed effects. Standard errors clustered by tract in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Finally, I estimate the inverse elasticity of labor demand ($\alpha - 1$) using demand shocks to nearby census tracts as an instrument. Results, shown in Table E2 vary between -0.23 and -0.33 when only employment is taken into account, implying labor's share of income is roughly 0.7. Column 2 of Table E2 includes the own-tract demand shock, Δz_{jt}, as a control (recall that the instrument is $\Delta z_{jt}^{LS}(\rho)$). This permits limited spatial correlation (to the extent the observed labor demand shocks are spatially correlated), and implies a slightly higher labor share of income. Column C includes the log measure of land zoned for productive uses, but this is measured poorly in the data.[E.1] Similarly, Column D indicates too large estimates.

E.1. Unlike residential land, it is difficult to classify different types of land used in production. For example, it is unclear whether to add land used for storage. Further, the data show some unusual changes across waves.

Table E2: IV estimates of inverse labor demand elasticity ($\alpha - 1$)

	$w_{jt} = \ln(W_{jt})$			
	(1)	(2)	(3)	(4)
ln(Employment)	-0.329**	-0.226**	-0.835	
	(0.125)	(0.082)	(0.698)	
ln(Prod. Land)			1.210	
			(0.980)	
ln(Emp. Density)				-0.553
				(0.368)
F-stat (KP)	3.586	2.955	1.798	3.439
Own shock as control	-	Y	Y	Y
N	4882	4882	4766	4766

Panel instrument variable (IV) estimates of regression of employment, employment density and land in production, using $\ln(\rho) = -5.5$. KP refers to the Kleinbergen-Papp F-statistic. Variables are trimmed to exclude extreme values (see text). Columns 2-4 include the own tract labor demand shock as a control. Standard errors clustered by tract in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Reasonable estimates of $\tilde{\alpha}$ and $\tilde{\zeta}$ provide confidence in this interconnected approach to identification and serve as an informal test of overidentification. Estimation of these parameters is more demanding than ϵ and ψ, both in terms of the stringency of the moment conditions and in the amount of exogenous variation needed to avoid weak instrument problems. Overall, these results suggest that interacting locally defined labor demand shocks with spatial structure can be used to create broad, omnipurpose tools for identifying local price elasticities.

E.1 Agglomeration in Productivity and Residential Amenity

To describe how the presence of agglomerative forces change identification assumption, define residential and productive spillovers as in Ahlfeldt et al. (2015):

$$\text{Productive agglomeration (A-augmenting):} \qquad \Upsilon_{jt} = \Upsilon\left(\sum_s k_{\Upsilon,js}\left(\frac{N^Y_{st}}{L^Y_{st}}\right)\right)$$

$$\text{Residential agglomeration (B-augmenting):} \qquad \Psi_{it} = \Psi\left(\sum_r k_{\Psi,ir}\left(\frac{N^H_{rt}}{L^H_{rt}}\right)\right)$$

where k here represents distance kernals and $N^H_i = \bar{N}\sum_s \pi_{is}$ is residential (employed) population.

If the parameters for the spillovers are known (of both the effects and the distance functions), then it is not necessary to develop new identification assumptions. Instead, the following substitutions can be made:

$$w_{jt} - \ln(\Upsilon_{jt}) \text{ for } w_{jt} \text{ in the labor demand equation}$$
$$\theta_{it} - \ln(\Psi_{it}) \text{ for } \theta_{it} \text{ in the housing demand equation}$$

Note that these equations reveal why the presence of these forces has little effect in this setting: They are mostly captured by the fixed effects $\bar{a}_j$ and $\bar{b}_i$.

If the spillovers are omitted from the model, additional moment conditions are required. Moment conditions presented in Assumptions A-1, A-1a, A-2, and A-2a do not change. Recall that those assumptions identify the key parameters of interest. Moment conditions corresponding to A-3, A-3a, and A-4 are tightened:

$$\mathbb{E}[\Delta z_{j't}^{LD,R} \times \Delta \ln(B_{it}\Psi_{it}D_{ijt})] = 0, \; \forall \; ij' \neq ij$$
$$\mathbb{E}[\Delta z_{j't}^{LD,R} \times \Delta \ln(B_{it}\Psi_{it})] = 0, \; \forall \; i$$
$$\mathbb{E}[\Delta z_{j't}^{LD,R} \times \Delta \ln(A_{jt}\Upsilon_{jt})] = 0, \; \forall \; j' \neq j$$

For these to hold, two additional assumptions are required in addition to Assumptions A-3 (or A-3a) and A-4:

$$\mathbb{E}[\Delta z_{j't}^{LD,R} \times \Delta \ln(\Psi_{it})] = 0, \; \forall \; i \tag{S-3}$$
$$\mathbb{E}[\Delta z_{j't}^{LD,R} \times \Delta \ln(\Upsilon_{jt})] = 0, \; \forall \; j' \neq j \tag{S-4}$$

If these conditions hold in addition to Assumptions A, the model is identified.

However, recall that instrument relevant requires $\mathbb{E}[\Delta z_{jt}^{LD,R} \times \Delta \ln(A_{jt})] \neq 0$. Both Ψ and Υ depend on nearby density, so to the extent location j' is near i or j, productivity shocks influence density and Assumptions S-3 and S-4 are unlikely to hold in a strict sense. However, they may hold approximately: There is significant autocorrelation in the population mass in locations from decade to decade. While this makes separately identifying agglomeration force difficult, in the context of the model presented here, this stickiness aids identification because much of $\Delta\Psi$ and $\Delta\Upsilon$ are captured by time-invariant tract fixed effects.

E.2 Endogenous Land Use

If land use is observed (as here) and the amount of land used in housing and production is determined by market forces, no additional assumptions need be made for identification. This is not true for the theoretical model or counterfactual simulations; both would need to be modified with an additional market clearing condition to account for the additional degree of freedom.

One minor change in interpretation of parameter values must be made if land use is endogenous. The assumption of congestion in the relationship between land price and residential density can no longer be supported: $P_i^L \neq (H_i/L_i^H)^{\tilde{\psi}}$. This is because of the price of land also depends on the demand for land for production (and so congestion occurs through displacing employment instead of density costs). ψ has no role in this alternate model. However, because total output (housing) is observable, we can modify the model to derive an estimating equation very similar to that in the main paper.

Consider the developer's problem. Zero profits implies $Q_iH_i = P_i^L L^H + P^M M$, and the first order conditions deliver an expression for M under profit maximization. This results in the expression:

$$Q_iH_i = \frac{1}{\phi}P_i^L L^H$$

which just requires that a constant fraction of developer income be spent on land. Solving this for P_i^L and substituting into Equation (6) and solving for Q_i delivers the equilibrium expression:

$$Q_i = \left(\frac{H_i}{L_i^H}\right)^{\frac{\phi}{1-\phi}} \mathfrak{C}_i$$

where $\mathfrak{C}_i = \frac{1-\phi}{\phi^2}P_M\tilde{C}_i^{1/(\phi-1)}$ contains the same elements as C_i. In fact, the estimating equation based on the above expression is isomorphic to that in the main text. Here, however, we identify $\frac{\phi}{1-\phi}$ instead of ψ. Note that under this interpretation, ϕ (the share of land in construction costs) is between 0.54 and 0.66. This is higher than a relatively standard value of 0.25 from Combes, Duranton, and Gobillon (2012), Epple, Gordon, and Sieg (2010), and Ahlfeldt et al. (2015). However, in Southern California land value anecdotally makes up high share of transacted real estate value. Alternatively, this could be seen as evidence in favor in immutable zoning.

As a quick aside, to complete the theoretical model, it is necessary to specify a land market clearing condition. I assume that the total land in a tract available for any use is fixed at $\bar{L}_i$; market clearing then requires $L_i^H + L_i^Y = \bar{L}_i$.[E.2] This condition can be rewritten (using Equation 4):

$$H_i \left(\frac{\mathfrak{C}_i}{Q_i} \right)^{\frac{1-\phi}{\phi}} + N_i^Y \left(\frac{W_i}{\alpha A_i} \right)^{\frac{1}{1-\alpha}} = \bar{L}_i$$

This equation, in conjunction with the model in the main text, is sufficient to pin down land use.[E.3]

E.3 Agglomeration and Endogenous Land Use

Because endogenous land use did not alter identification, identification with both agglomeration and endogenous land use requires the same assumptions as for the case with agglomeration: Assumptions S-3 and S-4 in addition to Assumptions A.

E.2. Note that this implies $\Delta L_{it}^Y = -\Delta L_{it}^H$.

E.3. Note that we can also rewrite this market clearing condition as an analytic expression of the observable prices, quantities, parameters, and the unobservable price of land:

$$\frac{\phi Q_i H_i}{P_i^L} + N_i^Y \left(\frac{(1-\alpha) W_i N_i^Y}{\alpha P_i^L} \right)^{\frac{1}{\alpha}} = \bar{L}_i$$

The price and land can be be calculated from this expression.

F Additional and Supplementary Results

Summary of Additional Results

- Figure F1: Ridership, 1990-2000
- Figure F2: Ridership, 1990-2014
- Figure F3: Glossary of variables and parameters
- Figure F4: Timeline of transportation in Los Angeles
- Table F1: Testing parallel pre-trends in other (non-commuting) tract characteristics, 1970-1990
- Table F2: Testing parallel pre-trends in tract-level commuting behavior, 1970-1990
- Table F3: Is treatment related to zero flows?
- Table F4: Effect of transit on commuting flows (by 2000) - PPML
- Table F5: Tracts near stations on the same line
- Table F6: Interactions of residential and workplace station proximity
- Table F7: Transit and non-commuting fundamentals (other effects of transit), robustness
- Table F8: Transit, income change, and land use change

Appendix Figures and Tables

Figure F1: Ridership, 1990-2000

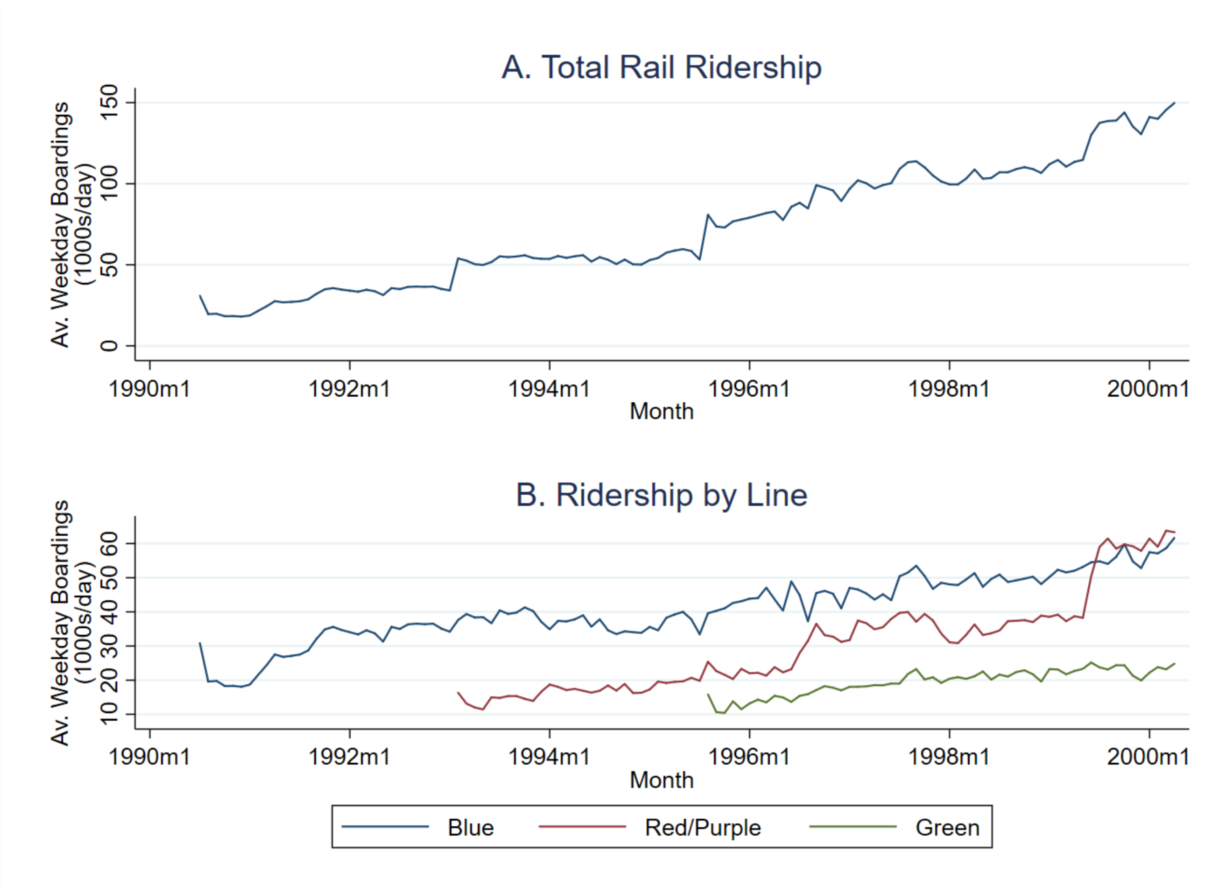


Figure F2: Ridership, 1990-2014

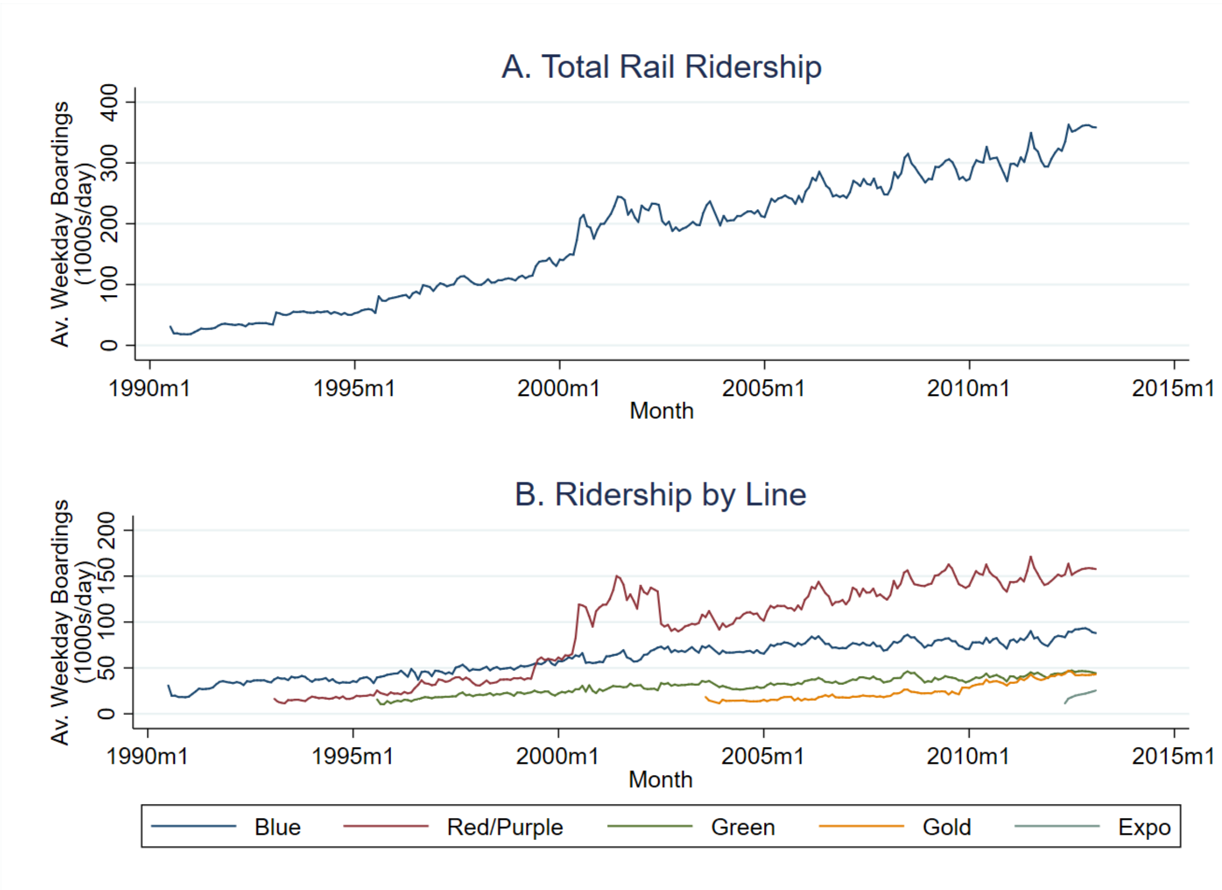


Figure F3: Glossary of variables and parameters

Parameters	Interpretation
ϵ	Homogeneity of location preferences (and wage elasticity of labor supply)
ζ	Household expenditure share on non-housing goods
$\tilde{\zeta} = -\epsilon(1-\zeta)$	Price elasticity of housing demand
α	Share of (production) income spent on labor
$\tilde{\alpha} = \alpha - 1$	Inverse wage elasticity of labor demand
ϕ	Share of housing income spent on land
$\tilde{\psi}$	Congestive cost of housing
$\psi = \tilde{\psi}\phi$	Inverse price elasticity of housing supply
κ	Semi-elasticity of commuting with respect to travel time
ρ	Spatial decay for instrumental variable construction
λ^x	Treatment effect for outcome x

Variables	Interpretation
A	Workplace productivity
$B = T\tilde{B}^{\epsilon}$	Gross residential amenity
$\tilde{B}$	Simple residential amenity
T	Mean residential utility
C	Inverse housing efficiency
$\tilde{C}$	Housing productivity
D	Mean utility commute (net of time)
E	Workplace amenity (net of wage)
$\mathcal{C}$	Consumption
H	Housing quantity
W	Wage
Q	Housing price
$\delta = e^{\kappa\tau}$	Commuting friction
τ	Travel time
π	Commuting share
$\bar{N}$	Aggregate population
N^Y	Employment at place of work
L^Y	Land used for production
M	Housing materials
L^H	Land used for housing
P^M	Price of housing materials
$P^L = (H/L^H)^{\psi}$	Price of land

Figure F4: Timeline of transportation in Los Angeles

1925	Comprehensive Rapid Transit Plan for the County of Los Angeles, Kelker, De Leuw & Co. developed at the request of local governments
1951	Los Angeles Metropolitan Transit Authority (LAMTA) formed
1961	Pacific Electric (Red Cars) end of service
1963	Los Angeles Railway (Yellow Cars) end of service
1964	Southern California Rapid Transit District (SCRTD) formed from LAMTA
3/24/1985	Ross Dress for Less methane explosion in Wilshire-Fairfax
1985	Construction begins on LA Metro Rail
11/20/1985	Department of Transportation and Related Agencies Appropriation Act (1986) includes language prohibiting funding of tunnels for transit along Wilshire corridor due to concerns about methane (HR 3244)
7/14/1990	Blue Line opens
2/15/1991	Metro Center station opens
1993	Los Angeles County Metropolitan Transportation Authority forms from SCRTD
1/30/1993	Red Line opens, connects system to Union Station
10/14/1993	Century Freeway (I-105) opens
8/12/1995	Green Line opens in median of Century Freeway
7/13/1996	Red Line expands to Wilshire/Vermont
6/12/1999	Red Line expands to Hollywood/Vine
6/24/2000	Red Line expands to North Hollywood
7/26/2003	Gold Line opens
2006	Purple Line renamed from Red Line branch
9/20/2006	HR 3244 amended to remove prohibitions on funding of tunnels for transit along Wilshire corridor
11/15/2009	Gold Line expands in East LA
4–6/2012	Expo Line opens
3/5/2016	Gold Line expands to Azusa
5/20/2016	Expo Line expands to Santa Monica

Table F1: Testing parallel pre-trends in non-commuting tract-level characteristics, 1970-1990

	ln Res. Emp. (1)	ln #HHs (2)	ln HHI (3)	ln House Value (4)
Subway Plan (Immediate) Sample				
Proximity$_i^{500m} \times t$	0.029	-0.011	-0.013	-0.002
	(0.020)	(0.017)	(0.013)	(0.019)
N	1629	1629	1628	1555
Subway Plan (All) Sample				
Proximity$_i^{500m} \times t$	0.012	-0.031^{+}	-0.019	-0.017
	(0.020)	(0.017)	(0.012)	(0.018)
N	3786	3786	3779	3688
PER Sample				
Proximity$_i^{500m} \times t$	0.002	-0.032^{+}	-0.020	-0.034^{+}
	(0.021)	(0.017)	(0.013)	(0.018)
N	4631	4629	4619	4502
Full Sample				
Proximity$_i^{500m} \times t$	0.025	-0.027	-0.016	-0.022
	(0.020)	(0.017)	(0.012)	(0.017)
N	11651	11641	11567	11407
Tract FE	Y	Y	Y	Y
Sbcty-X-Yr FE	Y	Y	Y	Y

Each column of each panel presents the results of a different regression, for 16 total. Estimates show pre-trends from 1970-1990 for tracts treated by 1999. All regressions include tract and subcounty-by-year fixed effects. Standard errors clustered by tract in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F2: Testing parallel pre-trends in tract-level commuting behavior, 1970-1990

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Commuting by automobile								
Proximity$_i^{500m} \times t$	-0.002	-0.001	-0.003	-0.001	0.000	-0.000	0.001	-0.002
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
N	11686	11644	1632	1629	3792	3786	4643	4631
B. No car households								
Proximity$_i^{500m} \times t$	-0.146**	-0.012	-0.006	0.014	-0.028	0.021	-0.044	0.007
	(0.035)	(0.036)	(0.040)	(0.040)	(0.055)	(0.038)	(0.055)	(0.037)
N	7720	7692	1086	1084	2524	2520	3086	3078
C. Transit (rail and bus) commuters, >0								
Proximity$_i^{500m} \times t$	-0.204**	0.023	0.043	0.042	-0.117**	-0.007	-0.135**	0.010
	(0.038)	(0.044)	(0.040)	(0.043)	(0.040)	(0.043)	(0.040)	(0.044)
N	9708	9669	1617	1614	3726	3721	4459	4448
D. Transit (rail and bus) commuters, all								
Proximity$_i^{500m} \times t$	-0.001	0.101**	0.098**	0.089**	0.051*	0.088**	0.022	0.097**
	(0.020)	(0.022)	(0.021)	(0.022)	(0.020)	(0.023)	(0.021)	(0.023)
N	11261	11195	1626	1623	3786	3776	4629	4617
Sample	All	All	Sim	Sim	Sal	Sal	PER	PER
Tract FE	Y	Y	Y	Y	Y	Y	Y	Y
Sbcty-X-Yr FE	-	Y	-	Y	-	Y	-	Y

Each column of each panel presents the results of a different regression, for 32 total. Estimates show pre-trends from 1970-1990 for tracts treated by 1999, except for Panel B, which only covers 1980-1990. Panel C is log-linear; Panels A, B, and D are estimated by PPML with exposure set to relevant tract population. All regressions include tract fixed effects. Tracts are 2010 geography. Standard errors clustered by tract in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F3: Is treatment related to zero flows?

	$1_{N_{ijt}>0}$ (1)	$1_{N_{ijt}>0}$ (2)	$1_{N_{ijt}>0}$ (3)
O & D <500m from station	0.032^{+}	0.020	0.021
	(0.018)	(0.016)	(0.016)
N	1260324	1259720	1259720
Control Network	All	All	All
Tract Pair FE	Y	Y	Y
POW-X-Yr FE	Y	Y	Y
RES-X-Yr FE	Y	Y	Y
Sbcty-X-Sbcty-X-Yr FE	-	Y	Y
Travel Time	-	-	Y

High-dimensional fixed effects estimates of transit on an indicator for positive flows. Control network is 'loose' (see text). All estimates include tract of work-by-year, tract of residence-by-year, and tract pair fixed effects. Travel time is measured in minutes. Standard errors clustered by tract pair, tract of residence, and tract of work in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F4: Effect of transit on commuting flows (by 2000) - PPML

	(1)	(2)	(3)	(4)	(5)	(6)
Subway Plan (Immediate) Sample						
O & D contain station	0.101	0.150^{+}	0.152^{+}	0.169^{*}	0.163^{*}	0.193^{*}
	(0.063)	(0.079)	(0.080)	(0.080)	(0.081)	(0.087)
O & D <250m from station			0.108	0.125	0.120	0.176^{*}
			(0.079)	(0.078)	(0.079)	(0.082)
O & D <500m from station		0.085	0.070	0.064	0.063	0.094
		(0.055)	(0.059)	(0.057)	(0.057)	(0.058)
N	69614	69614	69614	69596	69596	61922
Subway Plan (All) Sample						
O & D contain station	0.105	0.141^{+}	0.142^{+}	0.134^{*}	0.129^{+}	0.162^{*}
	(0.065)	(0.073)	(0.073)	(0.067)	(0.068)	(0.067)
O & D <250m from station			0.138^{*}	0.125^{+}	0.120^{+}	0.130^{+}
			(0.070)	(0.066)	(0.068)	(0.068)
O & D <500m from station		0.111^{*}	0.093	0.074	0.072	0.077
		(0.050)	(0.057)	(0.050)	(0.051)	(0.052)
N	309700	309700	309700	309700	309700	299472
PER Sample						
O & D contain station	0.089	0.126^{+}	0.127^{+}	0.110^{+}	0.106	0.054
	(0.065)	(0.072)	(0.072)	(0.066)	(0.068)	(0.069)
O & D <250m from station			0.149^{*}	0.133^{*}	0.128^{+}	0.058
			(0.068)	(0.065)	(0.067)	(0.067)
O & D <500m from station		0.125^{*}	0.110^{*}	0.087^{+}	0.085^{+}	0.020
		(0.049)	(0.056)	(0.049)	(0.050)	(0.049)
N	406494	406494	406494	406450	406450	383678
Full Sample						
O & D contain station	0.134^{+}	0.163^{*}	0.163^{*}	0.125^{+}	0.124^{+}	
	(0.069)	(0.074)	(0.074)	(0.065)	(0.065)	
O & D <250m from station			0.134^{*}	0.101^{+}	0.100	
			(0.067)	(0.061)	(0.063)	
O & D <500m from station		0.131^{**}	0.130^{*}	0.079	0.078	
		(0.049)	(0.058)	(0.048)	(0.049)	
N	1259500	1259500	1259500	1256986	1256986	
Control Network	Loose	Loose	Loose	Loose	Loose	Tight
Tract Pair FE	Y	Y	Y	Y	Y	Y
POW-X-Yr FE	Y	Y	Y	Y	Y	Y
RES-X-Yr FE	Y	Y	Y	Y	Y	Y
Sbcty-X-Sbcty-X-Yr FE	-	-	-	Y	Y	Y
Highway Control	-	-	-	-	Y	Y

High-dimensional fixed effects estimates of λ^D. Treatment variables are mutually exclusive with others in each column. All estimates include tract of work-by-year, tract of residence-by-year, and tract pair fixed effects. Standard errors clustered by tract pair, tract of residence, and tract of work in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F5: Tracts near stations on the same line

	(1)	(2)	(3)	(4)
O & D contain station, same line	0.205**	0.192**	0.153*	0.144*
	(0.077)	(0.064)	(0.062)	(0.059)
O & D contain station, not same line	0.075	0.089	0.058	0.042
	(0.091)	(0.079)	(0.078)	(0.076)
O & D <250m from station, same line	0.145*	0.112*	0.093+	0.062
	(0.066)	(0.055)	(0.054)	(0.051)
O & D <250m from station, not same line	0.105	0.093	0.085	0.047
	(0.078)	(0.068)	(0.067)	(0.065)
O & D <500m from station, same line	0.041	0.045	0.046	0.014
	(0.054)	(0.041)	(0.040)	(0.038)
O & D <500m from station, not same line	-0.048	-0.052	-0.037	-0.073
	(0.066)	(0.054)	(0.054)	(0.052)
Control Network	1925 Imm	1925 All	PER Lines	All
Tract Pair FE	Y	Y	Y	Y
POW-X-Yr FE	Y	Y	Y	Y
RES-X-Yr FE	Y	Y	Y	Y
Sbcty-X-Sbcty-X-Yr FE	Y	Y	Y	Y
Highway Control	Y	Y	Y	Y
N	19222	74040	99054	290580

High-dimensional fixed effects estimates of λ^D. Treatment variables are mutually exclusive with others in each column. All control networks are 'loose' (see text). All estimates include tract of work-by-year, tract of residence-by-year, and tract pair fixed effects. Standard errors clustered by tract pair, tract of residence, and tract of work in parentheses: $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F6: Interactions of residential and workplace station proximity

	D contains station	D<250m from station	D<500m from station
O contains station	0.140**	0.078	0.083
	(0.045)	(0.079)	(0.113)
O<250m from station	0.024	0.018	0.054
	(0.051)	(0.066)	(0.057)
O<500m from station	0.197*	-0.100	0.059
	(0.077)	(0.089)	(0.064)
Control Network	1925 Plan (All), Loose		
Tract Pair FE	Y		
POW-X-Yr FE	Y		
RES-X-Yr FE	Y		
Sbcty-X-Sbcty-X-Yr FE	Y		
Highway Control	Y		
N	74040		

High-dimensional fixed effects estimates of λ^D from one regression. Treatment variables are mutually exclusive all others. All estimates include tract of work-by-year, tract of residence-by-year, and tract pair fixed effects. Standard errors clustered by tract pair, tract of residence, and tract of work in parentheses: $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F7: Transit and non-commuting fundamentals (other effects of transit), robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Effect on productivity ΔA**,** $\alpha - 1 = -0.25$								
Proximity$_i^{500m} \times t$	-0.093**	0.009	-0.009	0.007	-0.036	0.005	-0.052^{+}	0.010
	(0.028)	(0.030)	(0.032)	(0.034)	(0.028)	(0.031)	(0.028)	(0.031)
N	4882	4858	780	776	1828	1826	2288	2284
B. Effect on residential amenity ΔB**,** $\epsilon(1-\zeta) = 0.125$								
Proximity$_i^{500m} \times t$	0.059*	-0.006	0.026	-0.029	0.018	-0.009	0.021	-0.002
	(0.024)	(0.026)	(0.027)	(0.029)	(0.024)	(0.027)	(0.024)	(0.027)
N	4534	4518	712	710	1700	1700	2094	2092
C. Effect on inverse housing efficiency ΔC**,** $\psi = 2.292$								
Proximity$_i^{500m} \times t$	0.073^{+}	-0.009	-0.103*	-0.058	0.034	-0.013	0.063	0.013
	(0.039)	(0.044)	(0.044)	(0.050)	(0.040)	(0.046)	(0.040)	(0.047)
N	4534	4526	712	712	1694	1694	2086	2084
D. Effect on workplace amenity ΔE**,** $\epsilon = 1$								
Proximity$_i^{500m} \times t$	-0.245**	-0.050	-0.095^{+}	-0.137*	-0.119*	-0.092	-0.131*	-0.100^{+}
	(0.051)	(0.054)	(0.057)	(0.062)	(0.052)	(0.057)	(0.051)	(0.055)
N	4866	4842	780	776	1830	1828	2286	2282
E. Effect on workplace amenity ΔE**,** $\epsilon = 0.498$								
Proximity$_i^{500m} \times t$	-0.271**	-0.045	-0.097^{+}	-0.127*	-0.128*	-0.085	-0.147**	-0.091^{+}
	(0.049)	(0.053)	(0.054)	(0.059)	(0.050)	(0.054)	(0.050)	(0.053)
N	4866	4842	780	776	1830	1828	2286	2282
Sample	All	All	Sim	Sim	Sal	Sal	PER	PER
Tract FE	Y	Y	Y	Y	Y	Y	Y	Y
Sbcty-X-Yr FE	-	Y	-	Y	-	Y	-	Y
Controls	-	Y	-	Y	-	Y	-	Y

Results from 40 regressions of transit proximity on estimated local fundamentals. All regressions include tract fixed effects. Controls include changes in highway proximity and 1990 levels of log household income, share of residents with at least a high school degree, and manufacturing employment. Standard errors clustered by tract in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F8: Transit, income change, and land use change

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Change in residential land								
Proximity$_i^{500m} \times t$	-0.016**	0.006**	0.001	0.001	0.001	0.001	-0.000	0.002
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	4948	4930	774	770	1840	1838	2306	2300
B. Change in household income								
Proximity$_i^{500m} \times t$	-0.015	-0.016	0.003	-0.019	-0.001	-0.005	-0.006	-0.005
	(0.015)	(0.017)	(0.017)	(0.018)	(0.016)	(0.017)	(0.016)	(0.017)
N	4954	4940	762	760	1824	1824	2280	2278
Sample	All	All	Sim	Sim	Sal	Sal	PER	PER
Tract FE	Y	Y	Y	Y	Y	Y	Y	Y
Sbcty-X-Yr FE	-	Y	-	Y	-	Y	-	Y
Controls	-	Y	-	Y	-	Y	-	Y

Results from 16 regressions of transit proximity on residential land and household income measures. All regressions include tract fixed effects. Controls include changes in highway proximity and 1990 levels of log household income, share of residents with at least a high school degree, and manufacturing employment. Standard errors clustered by tract in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$