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

Christopher Severen†
Federal Reserve Bank of Philadelphia

December 2018

Abstract

Using a panel of tract-level bilateral commuting flows, I estimate the causal effect of Los Angeles Metro Rail on commuting between connected locations. This unique data, in conjunction with a spatial general equilibrium model, isolates commuting benefits from other channels. A novel strategy interacts local innovations with intraurban geography to identify all model parameters (local housing and labor elasticities). Metro Rail connections increase commuting between locations containing (adjacent to) stations by 15% (10%), relative to control routes selected using proposed and historical rail networks. Other margins are not affected. Elasticity estimates suggest relatively inelastic mobility and housing supply. Metro Rail increases welfare $146 million annually by 2000, less than both operational subsidies and the annual cost of capital. 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, Paulina Oliva, and Wei You, as well as several anonymous referees. First draft was December 10, 2015.

†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 consider a broad range of interventions to mitigate the costs of distance and congestion. Subway and light rail systems are increasingly popular: Since the 1980s, Atlanta, Austin, Dallas, Denver, Houston, Los Angeles, Miami, Phoenix, Portland, and Seattle have all built systems. But rail transit is expensive, costing far more per mile than roads.[1] The benefits of urban rail are less certain in polycentric cities built around private automobiles. Do these benefits outweigh costs in automobile-dominated modern cities?

I study the effects of Los Angeles Metro Rail on commuting and welfare. Los Angeles is a large, car-oriented region that built an extensive rail network within a decade, making LA Metro Rail particularly relevant for other cities considering rail transit. I assemble unique data that include all census tract-to-census tract commuting flows and times in 1990 and 2000 and develop an identification strategy that exploits both the bilateral and panel aspects of the data to provide the first direct estimates of the effect of transit on commuting flows. The gravity-like estimating equation includes origin-year and destination-year fixed effects, allaying standard selection concerns due to non-random placement of transportation infrastructure.[2] Instead, 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 pairs of locations in which just one, or neither, receive treatment.

I use three complementary strategies to select a plausible counterfactual transit network and recover causal estimates of transit's effect of bilateral commuting. The first two strategies exploit historical maps of streetcar and proposed subway routes, and the third defines control pairs by adjacency to treated pairs and provides a lower bound on the effect size. The 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%. By measuring commuting flows (rather than travel time), these results implicitly capture very local effects of changes in congestion brought about by transit.

I describe and carefully estimate a quantitative economic geography model to translate the commuting effects to welfare and account for general equilibrium effects. The model casts the city as a collection of labor and housing markets (census tracts) connected by commuting, and generates the gravity equation used to estimate commuting effects.[3] However, the model requires census-tract scaled labor and housing market elasticities that have not been rigorously estimated

1. Light rail in the US typically costs $50-$150 million per mile, while subways in the US typically cost more that $500 million per mile. At grade highways typically cost about $10 million per mile.

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

3. I show how the addition of rich commuting data allows this urban economic geography model to be expressed by a system three linear (in log) equations. In this respect, the model is a within city analog of Roback (1982), but with commuting. Albouy and Lue (2015) study larger-scale regional patterns using compensating differentials.

before. I therefore develop a new strategy and assemble unique data to identify these parameters. I first develop a local implementation of a shift-share instrument that exploits tract-level variation in labor demand within the city. I instrument previously unused data on changes in average wage at tract of work with this local labor demand shock to estimate a local (tract-level) labor supply elasticity. This elasticity governs how responsive agents are to changes in prices, amenities, and commuting costs, and is essential to translate treatment effects to utility. 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.

I next interact the tract-specific labor demand shocks with the spatial configuration of the city to generate additional instruments that identify all remaining labor and housing market elasticities. I find that housing supply is inelastic and estimate reasonable values for other parameters. In addition to parameterizing the model, this step recovers time-varying, tract-specific fundamentals that correspond to non-commuting primitives that determine housing and labor supply and demand (such as productivity and amenities). I estimate the effect of transit on these local fundamentals with a difference-in-difference strategy; control tracts are selected using the historical streetcar and proposed subway route data. This provides an explicit test whether transportation infrastructure alters city structure through commuting or through other channels. The commuting effect dominates; impacts from non-commuting channels (e.g., amenities) appear minimal.

Transit increases the attractiveness of commuting between connected tracts, leading to substantial welfare gains. Preferred estimates show that by 2000, LA Metro Rail generates $109-$146 million in annual surplus, or $6-$8 for every ticketed ride. However, these benefits amount to no more than one-half the annualized cost of construction and net operating expenses, depending on the discount rate. I draw upon alternative data to test for further changes in commuting after another fifteen years. 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 in commuting 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 output, commuting.[4]

I also assess common methods and assumptions in the growing urban economic geography literature (e.g., Ahlfeldt et al. 2015; Donaldson 2018; Monte, Redding, and Rossi-Hansberg 2018; Tsivanidis 2018) with panel data for all tract-tract (bilateral) commuting flows and average workplace wage; such data are typically unobserved. I compare observed commuting flows to standard practice, which uses cross-sectional travel survey data to estimate the marginal disutility of travel time and then infer changes in commuting from computer-modeled changes in travel

4. One important caveat is that while I calculate the commuting effects over a twenty-five year window, I can only examine other channels between 1990 and 2000. Unmeasured benefits include increased mobility for non-commuters (those unemployed or not in the labor force) and environmental factors.

time. Flow data are prima facie preferable, as changes in bilateral commuting implicitly reflect changes in congestion and other factors that are difficult to measure or model (like reliability, cost, etc.) Panel flow data offer better measurement of the marginal disutility of travel time (gravity); estimates are smaller than in the cross-section, reflecting the importance of time-invariant, pair-specific characteristics in determining commuting. Commuting flows are also informative about the connectivity between locations. Compared to measures of "market access"—which describe potential connectivity—commuting based measures represent more heterogeneity and thus reflect a less smooth, and likely more realistic, urban geography. Finally, common practice recovers workplace wage by assuming that wage perfectly determines employment (conditional on residential geography and travel time). These recovered wages poorly explain observed wages, and identifying assumptions that rely on these recovered wages may be implausible. Using this data, I find a smaller Fréchet shape parameter (labor supply elasticity) than typical, reflecting greater idiosyncrasy in location choice.

Quantifying the effects of transit is challenging. A hedonic literature notes that commuting benefits of transit are capitalized into housing and land prices (Baum-Snow and Kahn 2000; McMillen and McDonald 2004). However, this approach primarily considers homeowner benefits, potentially excluding other commuters. Nor do hedonics easily isolate the commuting benefit; residential amenities (Chen and Whalley 2012; Kahn 2007) or disamenities (Bowes and Ihlanfeldt 2001) may play a role. Price spillovers are another concern, potentially violating the stable unit treatment value assumption (SUTVA).[5] Metropolitan-level analysis suggests that public transit expansion may enhance aggregate productivity and employment growth (Chatman and Noland 2014; Duranton and Turner 2012), but has at most a small effect on population growth (Gonzalez-Navarro and Turner 2018) and does little to reduce city-level commute times (Duranton and Turner 2011). Such aggregate analysis avoids many issues, but cannot capture local factors or study urban form.[6]

The particular research setting is of great interest: Metro Rail installed a relatively large rail network with forty-six stations on four lines by 2000.[7] The experience of Los Angeles is more informative for most cities considering rail-based mass transit than studies 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

5. Gibbons and Machin (2005) show that transit can displace housing demand elsewhere; Donaldson and Hornbeck (2016) discuss the importance of modeling general equilibrium when evaluating transportation infrastructure.

6. At the other end of the spectrum, transportation analyses often follow McFadden (1974) to study modal choice. These models require knowledge of the components of travel costs (travel time, reliability, complementary activities, etc.), agents' choice sets, and accurate measures of their value. However, these models typically do not endogenize agents' workplace and residence decisions, and empirical estimates vary widely. Summarizing many studies, Small and Verhoef (2007) notes value of time estimating varying from 35%-84% of wage rate for commuting, and from 20%-90% of wage rate for personal travel.

7. Metro Rail officially ran three lines during this period, but one line had two branches and was later separated to form two separate lines.

role of cities as centers of consumption in addition to production (Baum-Snow, Kahn, and Voith 2005; Billings 2011; 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.[8] Relative to this literature, I use new data to study a new outcome (commuting flows), develope a comprehensive identification strategy, and report credible estimates of welfare impacts.[9]

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 structural elasticities that parameterize the model. I then report estimates of these elasticities and the non-commuting effects of transit in Section 7. Section 8 describes commuting welfare estimation and results, and Section 9 discusses extensions. Section 10 concludes.

2 Setting: Commuting and Transit in Los Angeles

In the 1980s, commuting in Los Angeles was dominated by the automobile. Among the five US Metropolitan Statistical Areas (MSAs) with at least 5 million residents, Los Angeles area 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, in combination with complex geography prone to bottlenecks, meant that Los Angeles was consistently ranked the most congested urban area in the United States. The average trip in LA took one-third longer than the uncongested time, three times the national average (Schrank et al. 2015).

Concerns about congestion were not new. Los Angeles had long been more car-oriented than other cities in the United States, adopting automobiles in large numbers during the rapid growth of the 1920s.[10] By the 1960s, momentum for a transit solution was growing. After several failed referendums, Los Angeles passed Proposition A in 1980, which enabled a sales tax increase dedicated to transit. Relative to prior proposals, Proposition A only suggested vague corridors to allow later community involvement. The plan would eventually combine subway and light rail operations to create an interconnected urban rail transit system. Construction began in 1985.

A heavy subway line was to run west from downtown along Wilshire Boulevard, an important employment corridor. Construction of the heavy subway line was altered after unexpected methane leakage resulted in an explosion in a discount clothing retailer on Wilshire in 1985 (see

8. Schuetz (2015) shows little change in employment near new Metro Rail stations, and Schuetz, Giuliano, and Shin (2016) asks 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.

9. The use of historical and proposed routes is a mainstay of regional economics (e.g., Baum-Snow 2007), but these tools are only just beginning to be applied within cities (e.g., Heilmann 2018). The adjacency approach applies the intuition of Dube, Lester, and Reich (2010) to bilateral flow data.

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

Elkind 2014). As a consequence, the line was instead routed along a more northerly corridor, although a compromise spur was allowed to extend some of the way along Wilshire. Westward expansion of this spur (later named the Purple Line) ended four miles before the initially proposed terminus due to Congressional legislation. The first five stations opened in early 1993. The three stations along the compromise spur opened in 1996. Three more stations opened along the northern alignment (later named the Red Line) in 1999. Finally, in mid-2000 three additional stations were added that crossed under the Hollywood Hills, connecting Downtown Los Angeles with the San Fernando Valley.

The heavy subway line was integrated into a larger network of light rail lines. The Blue Line opened before the Red/Purple Line in mid-1990 (though construction delays meant it did not reach its urban termini until early 1991). It was aligned at grade along a previous streetcar right of way, and was therefore the quickest to construct. The east-west Green Line opened in 1995, passing through southern Los Angeles county, partially in the median of a new highway alignment. The Green Line was originally intended to connect the rail system with the international airport, but was realigned to the south after the Federal Aviation Administration raised concerns about adverse impacts on flight paths.

Though I primarily focus on the system by 2000, 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 first 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.[11] 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 in this paper is a census tract or tract pair using 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. For data from 2000, I collect data at the block group or tract level.

11. I provide limited additional results using more recent (up to 2015) data despite these challenges.

I overlay the 2000 geographies on 1990 tract definitions to spatially assign them to 1990 tracts. For block groups that map into multiple tracts, I weight block groups to tracts proportionally by area.

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 geographies to 1990 and construct a 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 this data to exclude implausible changes between 1990 and 2000 levels (see Appendix for discussion). More recent CTPP products do not include workplace wage, which is the primary data limitation that restricts my primary analysis to the 1990-2000 period.

Transit data and treatment; other data sources. I obtain shapefiles with location data on Metro Rail transit stations and lines from the Los Angeles County Metropolitan Transportation Authority (LACMTA) and combine this 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. I obtain a panel of spatial land use data from the Southern California Association of Governments (SCAG). I also draw extensively from Kelker, De Leuw & Company (1925), which includes maps of a proposed subway system and of former streetcar lines operated by the Pacific Electric Railroad (PER).

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 potentially shift residential or workplace characteristics in addition to travel costs. Simple regression of commuting flows on transit will not generally differentiate commuting effects from other margins—even if well identified. I discuss identification the commuting effect of transit, $\frac{\partial N}{\partial \tau}\frac{\partial \tau}{\partial T}$, in this section, 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, such as complementarities between certain residential and workplace areas. 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 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 (and flexibly controlled for) by the time invariant pair fixed effects (e.g., Baier and Bergstrand 2007).[12] Unlike other gravity-based approaches, 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 binary definitions of treatment:

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*: both tracts have some part within 250 meters of a transit station,

iii) *O & D <500m from station*: both tracts have some part within 500 meters of a transit station.

For ease of interpretation, these are always *mutually exclusive* so as to represent sequentially less access to transit. For example, tracts that satisfy "O & D <250m from station" are recoded to not satisfy "O & D contain station." The median tract is 1.38km^2, so this bin roughly corresponds to

12. 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 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 generally qualitatively and quantitatively similar (see the Appendix). This is because most pairs that are ever zero (in either 1990 or 2000) are zero in both years. Always zero pairs do not contribute to alternative estimators (panel Poisson models), and so persistent zeros in panel data are less problematic than in the cross section.

a catchment area of (0km,1.42km] from the station. Only stations open before the end of 1999 are used to define treatment.[13]

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. There are few observable, time-varying, pair specific covariates to include in x_{ijt}, but two are potentially important. First, I add controls for proximity to the Century Freeway, which opened in the mid-1990s and along which the Green Line runs (see Brinkman and Lin (2017) for local effects of highways). Travel time is also important, but I do not have a time varying measure of travel time for all pairs of locations. Pair-fixed effects ς_{ij} capture most of this variation. Further, travel time may respond to transit (i.e., a bad control), as I discuss below.

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 tract pairs 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 selective placement of transit to connect pairs of locations that would have experienced increased commuting anyway. I limit selection in route placement by three complementary methods. The first two use historical data on proposed subway lines and former streetcar routes to restrict the sample to treated pairs of tracts and untreated (control) pairs of tracts that could have plausibly received transit but did not. The third approach uses a bilateral adoption of spatial adjacencies.

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.[14] This document also shows Pacific Electric Railroad (PER) lines installed in 1925.

13. 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.

14. The transit system was to be run by Southern Pacific Railroad, which had a significant (and perhaps overlarge)

The PER, colloquially called Red Cars, was an at-grade railway system that served Los Angeles through 1961. LA Metro Rail has extensively used former PER rights of way. I define control pairs as those for which either the origin or destination tract (or both) did not receive treatment but are within 1km of: (i) the Kelker, De Leuw and Co. subway proposal, "1925 Plan Sample"; or (ii) PER lines, "PER Sample". Maps from Kelker, De Leuw and Company (1925) are shown in Figure 1; treated and control tracts are mapped in Figure 2 (pairs are difficult to map).

Selecting these pairs to serve as controls is supported by four pieces of evidence. First, both treated and selected control tracts tend to be near historical transit corridors. Rail transit requires and generates relatively linear corridors, as exhibited by both transit and control routes. Proximity to historical transit also generated persistent changes in urban form that are still visible today (Brooks and Lutz 2016). Second, many control tracts have since received, or will soon receive, transit stations as Metro Rail expansions. The staggered rollout of lines and stations was largely to due to political expediency. Dueling factions of local government (supporting heavy or light rail) both wanted to claim credit for opening the first line (Elkind 2014). These factors suggest similar counterfactual evolution of commuting patterns between treatment and control pairs.

Third, the westward expansion of LA Metro Rail was delayed by an unexpected geologic shock. Original routing of the Red Line was along Wilshire Boulevard to Fairfax Avenue. Methane seepage into a nearby Ross Dress-for-Less store exploded on March 24, 1985, leading to federal legislation restricting tunneling along Wilshire. This high density corridor contains substantial residential development and employment, and appeared in almost every transit plan drawn up from the 1920s until today, and both 1925 Plan and PER Samples select this corridor as a control group. Finally, some treated tracts are inconsequentially treated, and some control tracts are incidentally untreated. For example, the Blue Line was meant to connect downtown Los Angeles and downtown Long Beach. The route had to pass through south-central Los Angeles to accomplish this, treating intermediate locations. But stations were not built at high frequency along the track, leading to some incidentally untreated locations.[15] Control pairs exploit this quasi-random variation in where and when stations opened.

I also adapt the adjacency strategy of Dube, Lester, and Reich (2010) to flow data. Groups are defined as a treated pair, along with all adjacent pairs for which at least one tract is not treated. Thus, for each treated group ij there is a group g of $\#_i \times \#_j$ tract pairs, where $\#$ represents the number of adjacent tracts plus one (the treated tract). I then estimate:

$$y_{ijgt} = \lambda T_{ijt} + \varsigma_{ij} + \varsigma_{gt} + \varepsilon_{ijt} \tag{4}$$

Each group is permitted to have its own time trend, ς_{gt}. Tracts pairs are thus being compared to only doubly nearby tract pairs, making the parallel commuting trends assumption more credible.

influence on regional politics (Fogelson 1967).

15. It is generally suboptimal to locate stations too close to each other (Crampton 2000).

However, estimates from this adjacency approach are prone to attenuation if the spatial scale of the effect is mismeasured, so estimates provide a lower bound on the true effect size. To mitigate this attenuation, I expand the immediate adjacencies to slightly more distant tracts in some specifications.

4.2 Commuting Flow Estimates

Table 2 indicates that LA Metro Rail led to an average increase of 10%-16% in commuting between tracts nearest transit stations by 2000. 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; more distant tracts more distant do not change. Results are largest using the 1925 Subway Plan control groups, varying between 11% and 16%. Estimates are smaller using the PER control group, and smallest (but still significant) with the full sample. Subcounty pair-by-year fixed effects slightly boost most point estimates. 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 most error structures: tract pair, residential tract, and workplace tract.

Tract-tract commuting data before 1990 are unavailable, so I cannot provide direct evidence of parallel pre-trends.[16] However, I compare some aspects of average commuting from tract of residence using NCDB data.[17] Tests of parallel pre-trends from 1970-1990 show little evidence of differences in commuting by automobile and not owning a car (Appendix Table F1). There is some evidence that selected locations were already experiencing an increase transit (bus) usage, although this effect goes away when conditioning on positive transit usage. Together these results suggest that the combination of historical controls and regional fixed effects makes substantial progress toward identification, and that transit was targeted to locations that were experiencing a shift from no transit usage to positive transit usage. These tract level characteristics are absorbed in the fixed effects in Equation 2.

Adjacency based estimates, shown in Table 3, support the finding that the tract pairs nearest stations experienced increased commuting. Standard errors are clustered by tract pair, residential tract, and workplace tract; some tract pairs belong to multiple groups. Column 1 uses tracts pairs both adjacent to those within 500m of a station as the control; the effect of transit is smaller but significant for tracts nearest a station. Columns 2 and 3 add tracts further away (adjacent to tracts

16. Pairs connected by 2000 had somewhat higher commuting in 1990 than controls, however, I find that higher flows in 1990 are associated with lower commuting in 2000 (see Appendix). Together, these results suggest that results in the full sample are downward biased. As expected, Table 2 shows smaller effects in the full sample.

17. 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.

within 1km of a station), and show larger effects for tracts nearest a station. Comparing specifications suggests some attenuation of the effect over space, and do not support increased commuting between slight more distant tracts.

The rich commuting flow data permit modeling heterogeneity by origin and destination proximity and tract connection (reported in the Appendix). Interacting proximity measures for origin and destination tracts indicate a greater effect of proximity at the destination than the origin (though estimates are less precise). This suggests commuters are more comfortable traveling a larger distance to a rail station from home than from a rail station to work. Additional results show that only pairs connected by the same line show a significant increase in commuting. I do not evidence of heterogeneous treatment effects by distance between origin and destination.

4.3 Commuting Time Estimates

A common expectation of transit is that it will relieve automobile congestion. I use reported changes in reported travel times from the CTPP to test whether rail transit decreases congestion:

$$\tau_{ijt} = \omega_{jt} + \theta_{it} + \varsigma_{ij} + \lambda^{\tau,\text{2km}} 1_{ij \text{ within 2 km},t} + \lambda^{\tau,\text{4km}} 1_{ij \text{ 2 to 4 km},t} + \varepsilon_{ijt} \tag{5}$$

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 capture routes most likely impacted.

Table 4 shows results for three measures of travel time: average travel time across all modes, log average travel time across all modes, 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; these fixed effects absorb potentially useful variation.

These results are also information about the long run attenuation of the results in Anderson (2014), who finds that a temporary labor strike disrupting LA Metro Rail service in 2003 increased automobile congestion near transit lines. 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 congestion in the long. The commuting time results in Table 4 are roughly one-quarter to one-third those in Anderson (2014), suggesting substantial, but not complete attenuation of congestion benefits.[18]

18. These commuting time results could confound the flow results in Section 4.2. This is unlikely: interacted subcounty pair-by-year fixed effects eliminate the time effect but not the flow effect, estimates of the flow effect are robust to including travel time, and adjacency estimates confirm a large flow effect for locations nearest stations. Nonetheless, including these effects in counterfactual exercises allows bounding the effects of any confounding on welfare.

4.4 Comparison with Market Access Terms

I compare the direct, commuting based measures of the effect of transit shown above to those implied by the 'market access' approach increasingly used to evaluate transportation infrastructure.[19] Market access terms abstract from two aspects of urban commuting. First, origin-destination pair fixed effects control for and retain rich, persistent variation in commuting between locations. At an urban scale, many idiosyncrasies (e.g., schools, parks, job types) determine commuting patterns. These slowly-evolving patterns are not captured by market access terms. Second, market access terms weight changes in connectivity by characteristics of the destination (for example, by population). Increasing connectivity between two populous locations can therefore have a large impact on market access, even when there is zero commuting between them.

The costs of these omissions can be evaluated with panel commuting flow data. Define two measures that summarize the relative commuting effect of transit at tract of residence:

$$\Delta\Xi_i^{CF} = \frac{\sum_s N_{is}}{\sum_s (1-\lambda^D T_{is}) N_{is}} - 1\ , \quad \Delta\Xi_i^{MA} = \frac{\sum_s e^{-\tilde{\kappa}\tau_{is}} Y_s}{\sum_s e^{-\tilde{\kappa}\tau_{is}} (1-\lambda^D T_{is}) Y_s} - 1$$

Ξ_i^{CF} uses commuting flow data while Ξ_i^{MA} is an employment-weighted market access measure. Employment at j is captured by $Y_j = \sum_r N_{rj}$ and, by analogy to the trade literature, is meant to represent market size. I construct two variants of market access: $\Xi_i^{MA,0}$ is the traditional definition that assumes all locations are connected, while $\Xi_i^{MA,+}$ excludes locations that never have a positive flow in any period (as Ξ_i^{CF} does implicitly).

Table 5 compares each of the market access measures to $\Delta\Xi_i^{CF}$ at treated locations for several values of $\tilde{\kappa}$. There are three immediate takeaways. First, market access terms require knowledge of κ, and estimates are sensitive to the value chosen. Second, the mean impact is similar for Ξ_i^{CF} and $\Xi_i^{MA,+}$, but $\Xi_i^{MA,0}$ has a much smaller mean. Intuitively, $\Xi_i^{MA,0}$ includes many more destinations to with which there is no change in commuting, decreasing the average. Third, both market access terms underpredict the variation in how impacted by transit locations are. Figure 3 visually compares the three terms using $\kappa = 0.05$.

The consequences of these empirical results are notable. The disutility of commuting time, κ, is typically estimated in the cross-section. I show (in Appendix F) that panel estimation leads to substantially smaller values.[20] Cross-sectional estimation also ignores pair-specific, time-invariant determinants of commuting. Distance and travel time only explain about 20% of the variation in

19. Standard practice proceeds: (1) Use a computer-modeled transportation network to evaluate transit infrastructure's impact on travel time. (2) Estimate the marginal disutility of travel time from a cross-sectional commuting survey. (3) Infer changes in commuting by combining (1) and (2). (4) Calculate change in access by interacting (3) with a characteristic (e.g., population or total income) of a nearby location, then sum over all nearby locations. This approach is useful because travel times are much easier to model or scrape than to observe in situ.

20. I find standard values of κ in the cross-section. With a pair fixed effect, κ is almost 0, though this regression is particularly prone to attenuation from measurement error. Two-step estimation that first estimates pair fixed effects from the panel then regresses these fixed effects on travel time in the cross-section consistently gives a $\kappa = -0.02$.

these fixed effects. Clearly, other elements matter as much or more than distance or travel time.

Differences in the mean and variance of predicted effects matter as well. In this setting, mean effects are smaller using market access terms, suggesting that these terms consistently underpredict changes in accessibility. At the same time, smaller variation dampens the effect of treatment on the most impacted places, generally smoothing the modeled economic impact across space. An approach based on commuting flows, while more data intensive, leads to a more accurate portrait of commuting behavior and responses.

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 that rationalizes spatial interactions between housing and labor markets. The city consists of a collection of N locations, operationalized as census tracts, that each contain a labor market and a housing market. There is no restriction on where agents live and work conditional on being within the city, so agents choose the location pair that maximizes utility conditional on commuting costs. The model links local, observable equilibrium outcomes to unobservable economic fundamentals.

The model is similar to that of Ahlfeldt et al. (2015), with five differences: (i) origin-destination pairs are subject to idiosyncratic preferences, (ii) local housing supply can be exogenously shifted, (iii) the model can be rewritten as a log-linear system of equations, (iv) land use is exogenously determined between housing and productive uses, and (v) agglomeration and consumption externalities are excluded from the primary model. The first two differences are useful generalizations that rationalize variation in observed commuting flows. The latter two match the empirical setting, simplify exposition, and have little impact (there is little scope for land use adjustment and externalities are essentially time invariant and captured by tract fixed effects). I detail how relaxing these assumptions alter identification in Section 6; results are generally robust to their incorporation.

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 $\delta_{ij} \geq 1$. 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 $H_{ijo} = (1-\zeta)W_j/Q_i$.[21]

To map indirect utility to choice probabilities, assume $\nu_{ij}(o)$ 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 chooses residential location i and place of work j is:

$$\pi_{ij} = \frac{\tilde{\Lambda}_{ij}\left(\delta_{ij} Q_i^{1-\zeta}\right)^{-\epsilon} (\tilde{B}_i W_j)^{\epsilon}}{\sum_r \sum_s \tilde{\Lambda}_{rs}\left(\delta_{rs} Q_r^{1-\zeta}\right)^{-\epsilon} (\tilde{B}_r W_s)^{\epsilon}} \tag{6}$$

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

21. 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.

city is:

$$\mathbb{E}[U_{ijo}] = \Gamma\left(\frac{\epsilon - 1}{\epsilon}\right) \cdot \left[\sum_r \sum_s \tilde{\Lambda}_{rs} \left(\delta_{rs} Q_r^{1-\zeta}\right)^{-\epsilon} (\tilde{B}_r W_s)^{\epsilon}\right]^{1/\epsilon} \tag{7}$$

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.[22] Firms select competitively available labor N^Y and land L^Y inputs to maximize profits under constant returns to scale.[23] Production is multiplicatively separable in local productivity A_j and a technology that is identical across j:

$$Y = A_j F(N_j^Y, L_j^Y) \tag{8}$$

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.[24] 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:

$$W_j = \alpha A_j \left(\frac{L_j^Y}{N_j^Y}\right)^{1-\alpha} \tag{9}$$

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.[25] Convexity in

22. 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.

23. A number of recent papers that compare metropolitan outcomes permit heterogeneity in workforce productivity, generally by education level (Diamond 2016). Other work has focused on the interaction of skill and the distribution of economic activity within and across cities (Davis and Dingel 2014). I abstract away from this in an effort to focus on very local effects — in the empirical application, what is one metropolitan statistical area in other papers is here more than 2,500 unique locations.

24. 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).

25. 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 thirty-year window). Including land use measures does not greatly change identification or results. There is little evidence of differential land use near transit.

land pricing serves as a congestive force, driving up prices in desirable locations until agents look elsewhere. As is standard in the literature, I specify housing production to take Cobb-Douglas form: $H = (L^H)^\phi M^{1-\phi}\tilde{C}_i$.[26] Developers sell housing in location i in a competitive market at per 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 is 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:

$$Q_i = \frac{P^M}{(1-\phi)\tilde{C}_i}\left(\frac{M}{L^H}\right)^\phi \tag{10}$$

Substituting this into 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.[27] This parameter provides a congestive force to the model. Combining the expression for land price with Equation (10) and compressing notation relates housing supply, price, and land availability:

$$Q_i = C_i\left(\frac{H_i}{L_i^H}\right)^\psi \tag{11}$$

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.

Equilibrium characterization

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

$$N_i^Y = \sum_r \bar{N}\pi_{ri} \tag{12}$$

26. 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.

27. 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.

Commuting shares (6) determine employment in any location. With frictional commuting, workers benefit from residing near work locations. Given the assumptions on household 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:

$$H_i = (1-\zeta)\sum_s \bar{N}\pi_{is}\frac{W_s}{Q_i} \tag{13}$$

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 (Existence and uniqueness)**.** *Consider the equilibrium defined by equations (6), (9), (11), (12), and (13):*

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

$$\frac{2\epsilon(\epsilon+1)(1-\alpha)(1-\zeta)}{1+\epsilon(1-\alpha)} - 1 \leq \frac{1}{\psi} \tag{14}$$

Proof. See Appendix. □

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 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.[28]

Recovering fundamentals

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

28. In this setting, agglomeration does not alter the equilibrium uniqueness condition; see the Appendix.

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 $\Lambda_{ij} = B_i E_j D_{ij}$.[29] Local fundamentals $\mathbf{A}$, $\mathbf{C}$, and $\mathbf{\Lambda}$ can be expressed as unique functions of data and parameters:

Proposition 2 (Residual uniqueness). *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. This is important as persistent, difficult-to-measure characteristics can play an anchoring role in cities (Lee and Lin 2018).

All components of the model are expressed in the commuting flow (6), wage setting (9), and housing price (11) equations. Taking logs delivers a tractable linear system:

$$w_j = g_0 + (\alpha - 1)n_j^Y + \ln(A_j) \tag{15}$$

$$n_{ij} = g_1 + \epsilon w_j - \epsilon(1-\zeta)q_i - \epsilon\kappa\tau_{ijt} + \ln(B_i E_j D_{ij}) \tag{16}$$

$$q_i = g_2 + \psi h_i + \ln(C_i) \tag{17}$$

where $n_j^Y = \ln\left(\bar{N}\sum_r \pi_{rj}/L_j^Y\right)$ is log employment density, $h_i = \ln\left((1-\zeta)\bar{N}\sum_s \pi_{is} W_s / Q_i L_i^H\right)$ is log housing density, and g capture remaining constants.[30] This system is a within city analog of the Roback (1982) and Rosen (1979) framework (often used to evaluate amenities) with commuting. Local fundamentals are potentially functions of covariates ($A = A(X)$ and so on) such as transit proximity. Equation (16) provides structural interpretation of the components in Equation (2).

29. This mapping marks significant divergence from the framework in 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 any location or pair specific variation in housing supply $\mathbf{C}$ or commute utility $\mathbf{D}$. Note that the components of $\mathbf{\Lambda}$ are not uniquely identified from the data; I use statistical arguments to separate $\mathbf{B}$, $\mathbf{E}$, and $\mathbf{D}$.

30. That is, $g_0 = \ln(\alpha)$, $g_1 = \ln(\bar{N}) - \ln\left(\sum_r \sum_s \Lambda_{rs}\left(e^{\kappa\tau_{rs}} Q_r^{1-\zeta}\right)^{-\epsilon} W_s^{\epsilon}\right)$, and $g_2 = 0$.

This system can be re-expressed to more clearly represent the supply and demand linkages. First, separate the unobservables into time varying and time invariant components, so that $\ln(A_{jt}) = \bar{a}_j + a_{jt}$, etc. This leads to the following system (omitting constants):

$$\begin{aligned}
\text{Labor demand in } i\text{:} \quad & w_{jt} = \tilde{\alpha} n_{jt}^{Y} + \bar{a}_j + a_{jt} && (18)\\
\text{Labor supply to } i\text{:} \quad & \omega_{jt} = \epsilon w_{jt} + \bar{e}_j + e_{jt} && (19)\\
\text{Commuting between } i \text{ and } j\text{:} \quad & n_{ijt} = \omega_{jt} + \theta_{it} - \epsilon\kappa\tau_{ijt} + \bar{d}_{ij} + d_{ijt} && (20)\\
\text{Housing demand in } i \quad & \theta_{it} = \tilde{\zeta} q_{it} + \bar{b}_i + b_{it} && (21)\\
\text{Housing supply in } i\text{:} \quad & q_{it} = \psi h_{it} + \bar{c}_i + c_{et} && (22)
\end{aligned}$$

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.

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 (18) to (22). 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. The moment conditions can identify all four elasticities, though I focus primarily on ϵ and ψ, as these two embed information about the local economic environment and cannot be estimated from microdata.[31] I also discuss the robustness of this strategy to alternative assumptions.

Summary

Identification requires a demand or supply shock that shifts one of Equations (15) to (17) 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.

31. In contrast, α and ζ could be estimated from microdata. Nonetheless, they are identifiable under some assumptions; estimates are reasonable.

This derived housing demand instrument traces out the housing supply curve.

Identifying housing demand requires an instrument that shifts housing supply. For agents who work in j and live in i, a labor demand shock to agents who work elsewhere (in j') but live in i 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} (i.e., 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. In the absence of spillovers in labor demand, the labor demand shocks in one location shift labor supply in nearby locations (conditional on the local shock), tracing out labor demand.

Detailed description of moment condition construction

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:

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

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. To ensure that local innovations in productivity do not drive national changes, I exclude all workers in California.

The demand shock instruments the change in wage to identify the slope of labor supply. Place of residence-by-year fixed effects control for changes in residential amenities that may be correlated with labor demand shocks (see Equation 16). The corresponding moment condition is:

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

Identification requires that the labor demand shock is uncorrelated with unobservable changes in labor supply (i.e., changes in workplace amenities, e, or travel costs, d). In fact, this can be weakened further by estimating place of work-by-year fixed effects to use as the dependent variable measuring labor supply conditional on commuting. I discuss and interpret this and the following identification assumptions later in this section.

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 $\Delta z_{it}^{HD,X} = \mathbf{z}_t^{LD,X} \cdot \boldsymbol{\vartheta}_i$. Weights ϑ are treated as parametric decay functions of the modeled travel time between locations that ever have positive commuting:

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

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. The spatial weight, $e^{-\rho\delta_{is}} 1_{\check{n}_{is}>0}$, means that labor demand shocks nearer a residential location with some commuting connection are more important than labor demand shocks farther away or in places with no commuting connection. The resulting inverse-distance weighted labor demand shock identifies ψ, the inverse price elasticity of housing supply:

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

Although both elements of M-2 relate to tract i, the housing demand shock draws on labor demand shocks from any j; I consider later how selecting particular subsets of these j may be useful.

Residents of a one location commute to 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 excluding own tract j:

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

Using the place of work-by-year fixed effects in Equation (16) to control for changes in workplace amenities, the following moment condition identifies $\epsilon(1-\zeta)$:

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

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 generate more variation than do individual tracts, this approach (along with an estimate of ϵ) recovers reasonable estimates of the household expenditure share on housing $1-\zeta$.

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 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:

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

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

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

This identifies the share of production income that goes to non-labor expenses, $\alpha - 1$, 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):

$$\mathbb{E}[\Delta z_{jt}^{LD,R} \times (\Delta e_{jt} + \Delta d_{ijt})] = 0, \ \forall \, ij \tag{A-1}$$

$$\mathbb{E}[\Delta z_{jt}^{LD,R} \times \Delta c_{it}] = 0, \ \forall \, ij \tag{A-2}$$

$$\mathbb{E}[\Delta z_{j't}^{LD,R} \times (\Delta b_{it} + \Delta d_{ijt})] = 0, \ \forall \, ij' \neq ij \tag{A-3}$$

$$\mathbb{E}[\Delta z_{j't}^{LD,R} \times \Delta a_{jt}] = 0, \ \forall \, j' \neq j \tag{A-4}$$

Proposition 3. *Assume A1, A2, A3, and A4 are true, $\rho > 0$, $\mathbb{E}[\Delta z_{jt}^{LD,R} \times \Delta w_{jt}] \neq 0$, 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, the presence of data on wages at place of work and commuting flows in combination with Equation (20) suggests high-dimensional fixed effects may be useful to control for unobserved confounders. Assumptions A-1 and A-3 can be weakened to exploit this:

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

Discussion and comparison to existing approaches

Assumptions A-1 and A-1a require that the local labor demand shocks be uncorrelated with changes in the local work amenity at j. They are weaker than standard identifying assumptions using labor demand shocks and than those common in the economic geography literature. When an aggregate labor demand shock is used to trace out 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, such a condition is

$$\mathbb{E}[\Delta f(\mathbf{z}_t^{LD,R}) \cdot \Delta f(\mathbf{B}_t, \boldsymbol{\delta}_t, \mathbf{D}_t, \mathbf{E}_t)] = 0$$

where f averages over locations. In contrast, Assumptions A-1 and A-1a clarify the spatial requirements for identification and are robust to local correlation between improvements in residential amenities and the productivity shocks.

The economic geography literature typically identifies ϵ from cross-sectional variation related to commuting. For example, Ahlfeldt et al. (2015) and Allen, Arkolakis, and Li (2015) condition on the time use parameter (κ) estimated from auxiliary models, and require that unobserved, origin-destination specific amenities are orthogonal to travel time. This is problematic because κ may be smaller than typically assumed (see Section 4 and Appendix). Ahlfeldt et al. (2015) also require that there be no variation in (non-pecuniary) workplace utility ($\mathbb{E}[\ln(E_j)^2] = 0$). I later show this assumption is improbable. Monte, Redding, and Rossi-Hansberg (2018) specify production within a trade framework and recover productivity from cross-sectional trade flows. They assume the productivity is orthogonal to workplace and origin-destination specific amenities.[32] Such indirect strategies are necessary because wage at place of work is typically unobserved. By comparison, Assumptions A-1 and A-1a utilize workplace wage data and the census of commuting flows, as well as rely on panel variation to control potentially confounding, persistent characteristics $\bar{e}_j$.

Assumption A-2 requires labor demand shocks be uncorrelated with changes in (inverse) housing productivity, Δc_{it}, which measures how efficiently developers provide housing density.

32. Using my notation, Ahlfeldt et al. (2015) first estimate $\epsilon\kappa$, and then use the assumption that $\mathbb{E}[\ln(E_j)^2] = 0$ to match the variation in modeled wages to highly aggregated observable wages. Tsivanidis (2018) estimates κ from models of commuting, and runs additional models instrumenting τ with plausibly exogenous variation in transit connections to limit endogeneity concerns. Tsivanidis (2018) also assumes $\mathbb{E}[\ln(E_j)^2] = 0$, but it does not impact identification of ϵ in his framework. Allen, Arkolakis, and Li (2015) estimate $\epsilon\kappa$, but then put more structure around time use to parameterize κ and identify ϵ. Monte, Redding, and Rossi-Hansberg (2018) assume $\mathbb{E}[\ln(A_j(\sigma)) \times \ln(E_j D_{ij})|\sigma, \kappa] = 0$, where σ comes from the trade model.

If the labor demand shocks are correlated with these housing productivity innovations, ψ is not identified. 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.

$$\mathbb{E}[\Delta z_{jt}^{LD,R} \times \Delta c_{it}] = 0, \; \forall \, i \neq j \tag{A-2a}$$

This 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; Saiz 2010), though the use of spatially heterogeneous labor demand shocks within cities to identify a local elasticity is novel.

Assumptions A-3 and A-4 are less central, as the parameters they identify ($\tilde{\alpha}$ and $\tilde{\zeta}$) can be estimated from microdata. Nonetheless, they can be viewed as providing an additional test of identification. Assumption A-3 requires that labor demand shocks in one location do not change amenities and commute utility in *other* locations, and controls for persistent residential and commuting amenities, $\bar{b}_i$ and $\bar{d}_{ij}$. That is, any innovations to amenities in j or the unobserved commute utility between i and j must be uncorrelated with national innovations in labor demand and the share of industry in 1990 in a third location $j' \neq j$. A potential cause for concern is that a large labor demand shock in location j' could change civic investment in public amenities in nearby locations i due to fiscal averaging within cities. Assumption A-3a is similar to A-3, but weaker in that the unobserved commuting shock does not enter. Assumption A-4 requires that labor demand shocks in one location be uncorrelated with nearby changes in productivity. Though tract fixed effects $\bar{a}_j$ control for most spatial correlation in industrial location, this identification assumption may not strictly hold. Nonetheless, estimates appear reasonable.

Identification with Agglomeration and Endogenous Land Use

This model abstracts away from agglomerative forces and endogenous land use determination. 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 residential amenities, though $\tilde{\alpha}$ and $\tilde{\zeta}$ can no longer be identified without ex ante knowledge of the parameters that govern these forces. Agglomerative forces only confound demand elasticities; given the exogenous demand shocks, supply elasticities are still identified. Agglomeration tends to be highly path dependent, therefore fixed effects $\bar{a}_j$ and $\bar{b}_i$ control for most of these forces (Davis and Weinstein 2008). Furthermore, Ahlfeldt et al. (2015) show that these forces mostly dissipate within a few (five) minutes of travel time, limiting their role to confound.

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). Assumptions A-2 and A-4 require productivity shocks to be orthogonal to changes in a location's ability to provide employment density and housing density. In other words, because these conditions are framed in terms of observable density by use, endogenous land use determination does not pose a direct threat to identification.[33]

6.2 Estimating the effects of interventions

Estimates of $\{\tilde{\alpha}, \epsilon, \tilde{\zeta}, \psi\}$ permit recovery of local economic fundamentals by removing the simultaneous components of the supply, demand, and commuting equations. 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 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:

$$n_{ijt} = \omega_{jt} + \theta_{it} - \epsilon\kappa\tau_{ijt} + \varsigma_{ij}^{D} + \lambda^{D} T_{ijt} + \varepsilon_{ijt}^{D} \tag{23}$$

$$\hat{\mathbf{Y}}_{it} = \boldsymbol{\lambda} T_{it} + \boldsymbol{\varsigma}_i + \boldsymbol{\varepsilon}_{it} \tag{24}$$

where $\boldsymbol{\lambda} = \{\lambda^A, \lambda^B, \lambda^C, \lambda^E\}$ and λ^D are the effects to be estimated. $\hat{\mathbf{Y}} = \{\widehat{\mathbf{a}}, \widehat{\mathbf{b}}, \widehat{\mathbf{c}}, \widehat{\mathbf{e}}\}$ contains the four non-commuting fundamentals, and Equation 23 corresponds to Equation 2 from the commuting analysis. 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.[34]

7 Housing and labor elasticities and non-commuting effects of transit

The shape parameter ϵ corresponds to homogeneity in location preference and represents an elasticity of labor supply that conditions on commuting and residential geography. I first re-

33. Endogenizing land use is not trivial, however. It requires an additional market clearing condition, changes properties of the model equilibrium, and alters counterfactual simulations. Such changes make a small numerical differences due to the constancy of land use in this context. Note that if land use were unobserved, identification would become more challenging as land use becomes a latent term.

34. 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 $\boldsymbol{\lambda} = \{\lambda^A, \lambda^B, \lambda^C, \lambda^E\} = 0$.

cover workplace by year fixed effects ω_{it} from the Equation 20 (ω_{it} has structural interpretation $\omega_{it} = \epsilon w_{it} + e_{it}$). I then estimate ϵ in differences using the wage variant of the labor demand shock as an instrumental variable (i.e., under M-1). Table 6 shows results using three different methods of estimating ω_{it}. Column 1 estimates ω_{it} from a linear panel, Column 2 uses a separate PPML estimator in each year with a measure of bilateral travel costs, and Column 3 uses a panel PPML estimator with pair fixed effects to control for travel costs. 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.[35] I take $\epsilon = 1.83$ from Column 3 as the preferred estimate. The low value of ϵ (relative to the literature) implies workers are quite heterogeneous in their location preferences.

Identification of ϵ in the economic geography literature sometimes requires $\omega_{it} = \epsilon w_{it}$. Such an assumption is not supported in my data (see Panel A of Figure 4).[36] Instead, ω_{it} is more closely related to workplace employment levels than wage (Panel B of Figure 4), highlighting the severity of the simultaneity problem. It is unlikely that that the ad hoc moment restrictions used in the literature correctly identify this parameter. The estimates presented here are closer to more standard estimates of labor supply elasticities (e.g., Falch 2010; Suárez Serrato and Zidar 2016). This low value has important implications for studies of urban structure, as preference heterogeneity limits the locational responsiveness of agents to changing local conditions.

The remaining structural parameters are identified using instruments constructed from the labor demand shocks and a spatial decay parameter, $\rho > 0$, that governs how the labor demand shocks propagate across space. I experiment with different values in $\ln(\rho) \in [-10, -2]$.[37] Labor demand shocks should propagate through the economy 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 $\ln(\rho) \in [-7.5, -4.5]$ based on gravity estimates (see Appendix). I report results for $ln(\rho) = -5.5$.

Estimates of the inverse housing supply elasticity (Equation 17) appear in Table 7. 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.[38] Column 2, 4, and 6 exclude

35. 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.

36. I run a regression of w_{it} on ω_{it}; the R^2 is only about 0.004. If $\omega_{it} = \epsilon w_{it}$, then $R^2 = 1$. Kreindler and Miyauchi (2017) find only a modest cross-sectional relationship between ω and w using cell phone data and travel surveys from Dhaka and Colombo.

37. 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.

38. If the assumption of Cobb-Douglas housing production is correct, then the coefficient on land should be negative and equivalent in magnitude to the housing level coefficient, ψ. The coefficients are not statistically different in absolute value from each other (Columns 3 and 4).

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).

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 8 and 9 roughly concur with values derived from other sources, lending additional credibility to the identification strategy as a whole. Table 8 gives estimates of Equation (16) using the employment variant of $\Delta z^{HS,X}_{i(-j)t}$ 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. Using $\epsilon = 1.83$, these imply a housing expenditure share between 36% and 48% of income, somewhat higher than microdata suggest but not unreasonable for high cost areas.[39]

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 9 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 9 includes the own-tract demand shock, $\Delta z^{LD,R}_{jt}$, as a control (recall that the instrument is $\Delta z^{LS,R}_{jt}(\rho)$). 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.[40] Similarly, Column D indicates too large estimates.

The ability to generate reasonable estimates of α and ζ provides confidence in this interconnected approach to identification. 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, omni-purpose tools for identifying local price elasticities.

39. IPUMS microdata indicate that the median household expenditure share on *renting* is about 0.26 for this time period, though there are a number of differences in calculating income and housing costs that could explain the mean difference.

40. 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.

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

I now test whether transit shifts non-commuting fundamentals using Equation (24). For tract-level analysis, I define transit proximity as

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

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 I compare pre-trends on observed market outcomes using NCDB data. The first four columns of Appendix Table F15 shows that there do not appear to be consistent differences in treated and control locations among modeled prices and quantities across specifications. I do, however, some evidence of different trends in socioeconomic characteristics. I include the 1990 levels of these variables as control in some specification to permit 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 10. Transit has little non-commuting effect after conditioning on regional trends and changes in highway structure. There some evidence of decreasing workplace amenities near transit, though an effect on this margin is unexpected. While this might represent an effect of transit on non-pecuniary workplace benefits (e.g. reduced parking), this is more likely due to labor match than a direct consequence of transit proximity. Taken together, these results indicate that transit is not generating large-scale non-commuting benefits or costs within the immediate proximity of stations.

Surprisingly, there is no apparent effect of LA Metro Rail on (non-commuting) residential amenities in any specification. An implication is that hedonic estimates of the effect of transit reflect a commuting benefit rather than other related neighborhood amenities.[41] These results only apply to LA Metro between 1990 and 2000; I cannot extend the non-commuting analysis to more recent years. The network was limited in size and connectivity at this time. As the transit network has expanded, it has become more valuable in terms of transportation connectivity. Responses that depend on scale, or are slower to respond (zoning laws can take decades to evolve), 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.

41. Results 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. Consumption amenities are a common foundation for residential agglomeration (Ahlfeldt et al. 2015).

7.2 Robustness checks: sorting and land use

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 F17 indicates the change in residential land use is very small and statistically insignificant across richer fixed effects specifications. This is unsurprising given strict zoning and the relatively fixed nature of land use in urban settings due to the slow depreciation of buildings, and accords with the finding in Schuetz, Giuliano, and Shin (2016) that zoning hinders transit-oriented development near rail stations.

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. Results in Table F17 show small, insignificant point estimates. Figure 5 shows the relationship between transit and rail usage by income centiles in 1990 and 2000, and reveals no relationship between income and rail usage.

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 under the counterfactual X'. This approach avoids using levels of 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.

8.1 LA Metro Rail and welfare

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 (7):

$$\%\Delta \text{ Welfare } \approx \ln \hat{\bar{U}} = \frac{1}{\epsilon} \ln \left(\frac{\hat{B}_i \hat{E}_j \hat{D}_{ij} \hat{W}_j^{*\epsilon} \hat{Q}_i^{*-\epsilon(1-\zeta)}}{\hat{\pi}_{ij}^*} \right) \tag{25}$$

for any pair ij, where $\hat{X}^*$ indicates the equilibrium value of X in the counterfactual under autarky (that is, fixing $\hat{\bar{N}} = 1$).[42] 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 2001/2 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 thirty 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 to $797 million per year (details in the Appendix).

Welfare effects by 2000

Table 11 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 $\zeta = 0.65$; $\psi = 1.693$ in Columns 1 and 2, and $\psi = 2.29$ in Columns 3 and 4.

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

42. On a technical note: the expectation underlying this term is infinite for $\epsilon \leq 1$, even though Equation (25) can be calculated for any value of $\epsilon \neq 0$. Instead of enforcing this arbitrary parameter restriction, I show (in the Appendix) that Equation (25) can be isopmorphically expressed in a multinomial logit framework, with ϵ naturally taking the roll of marginal utility of income (Train 2009). This expands permissible values of ϵ; Equation (25) (and similar expressions) can be used for $\epsilon > 0$, not just $\epsilon > 1$.

with Metro Rail. Including travel time increases the closed economy benefit to about $146 million per year, or about 0.07%. Open economy employment would be about 0.15% higher in this case (roughly ten thousand people). Closed economy results are virtually unchanged with less elastic housing supply, while open economy results are a bit smaller.

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 cleanly outweigh benefits.

Other margins

There are of margins to which the data and framework used for this analysis cannot speak. The framework does not capture or calculate the benefits for non-workers (as in most analysis of transportation behavior). This margin is likely quite important, and unfortunately understudied. Nor can I directly speak to benefits resulting in better transit provision for non-commuting trips, though this margin would likely show up as a residential amenity if substantial enough. Rail transit may also enable better bus transit and connectivity, but I cannot measure this. Finally, city-wide effects are not captured by this approach. For example, decreased air pollution may provide an additional benefit; a generous estimate using parameters from Gendron-Carrier et al. (2018) and Los Angeles' mid-1990s birthrate suggests an additional gain of about $180 million annually.[43]

9 Additional commuting effects

I use data from the 2002 and 2015 LEHD Origin-Destination Employment Statistics (LODES) to look for additional effects of transit on commuting flows in more recent years.[44] 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* is the effect of new stations (built after 2002) on bilateral commuting flows, while (ii) *Existing Transit* is the additional increase in commuting experienced by the stations built earlier (between 1990 and 2002). I retain prior treatment definitions.

Results (shown in Table 12) indicate that new transit connections increase commuting by 10%-13% between tract pairs that both contained stations by 2015.[45] 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

43. 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.

44. The 2006/10 CTPP is not compatible with previous CTPPs because of aggressive changes in reporting standards, aggregation of data across years, and other changes in geographies. Notably, it does not report workplace wage.

45. Note that this includes tract pairs for which one member had been connected before 2000.

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 either (i) aggregate commuting flows take decades to adjust to new transit modes (i.e., *habituation*), and/or (ii) that 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 12 can be included in longer run welfare analysis if I assume that, as was the case between 1990 and 2000, no non-commuting benefits accrue post 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 cost basis.

Under habituation, I simply combine the effects from Table 2 and the Existing Station effects from Table 12 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 11). 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 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 locations. Nearby stations 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 about 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 because of 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 ten to twenty-five years.

References

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

Albouy, David, and Bert Lue. 2015. "Driving to opportunity: Local rents, wages, commuting, and sub-metropolitan quality of life." *Journal of Urban Economics* 89:74–92.

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.

Brinkman, Jeffrey, and Jeffrey Lin. 2017. "Freeway Revolts!"

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.

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

Crampton, Graham R. 2000. "Urban economic structure and the optimal rail system." *Urban Studies* 37 (3): 623–632.

Davis, Donald R, and Jonathan I Dingel. 2014. "The comparative advantage of cities." *NBER Working Paper* w20602.

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. 2018. "Railroads of the Raj: Estimating the impact of transportation infrastructure." *American Economic Review* 108 (4-5): 899–934.

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

Dube, Arindrajit, T William Lester, and Michael Reich. 2010. "Minimum wage effects across state borders: Estimates using contiguous counties." *Review of Economics and Statistics* 92 (4): 945–964.

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.

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

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

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.

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."

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 US decennial census tract data from as early as 1970 to 2010: A longitudinal tract database." *Professional Geographer* 66 (3): 412–420.

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

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

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." In *Handbook of Regional and Urban Economics,* edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange, 5:1339–1398. 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 (114): 2699–2723.

Schuetz, Jenny, Genevieve Giuliano, and Eun Jin Shin. 2016. "Does Zoning Help or Hinder Transit Oriented (Re)Development?"

Small, Kenneth, and Erik Verhoef. 2007. *The Economics of Urban Transportation.* New York, NY: Routledge.

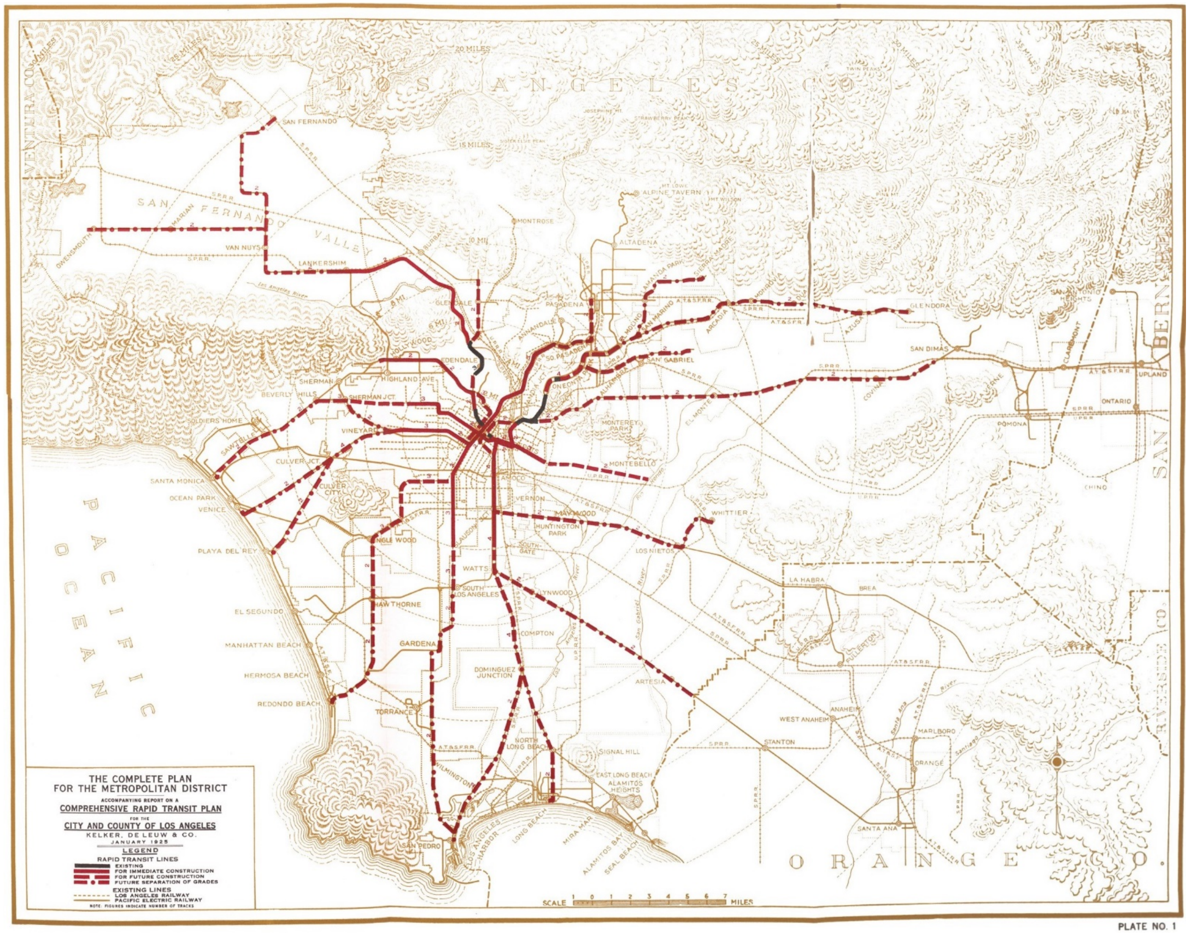
Suárez Serrato, Juan Carlos, and Owen Zidar. 2016. "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms." *American Economic Review* 106 (9): 2582–2624.

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 and Co. (1925)



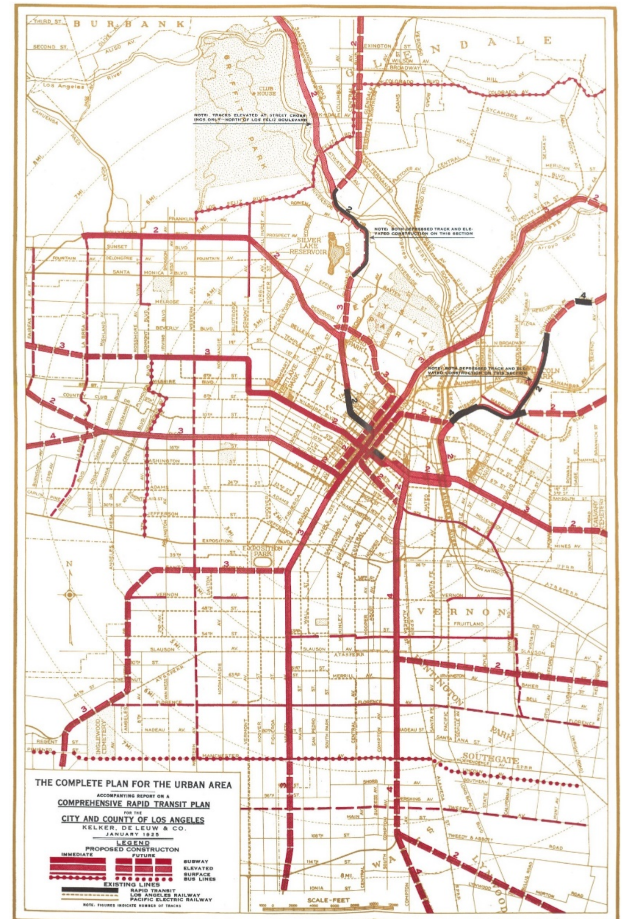
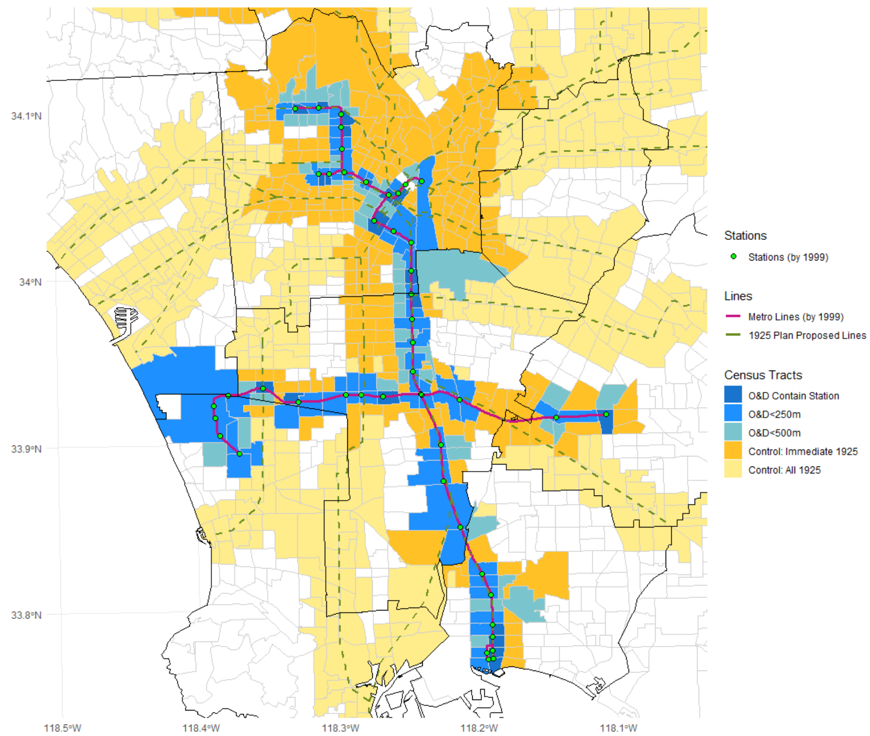


Figure 2: Map of LA Metro lines, stations, and the 1925 Plan and PER Lines

(a) 1925 Plan Sample



(b) PER Lines Sample

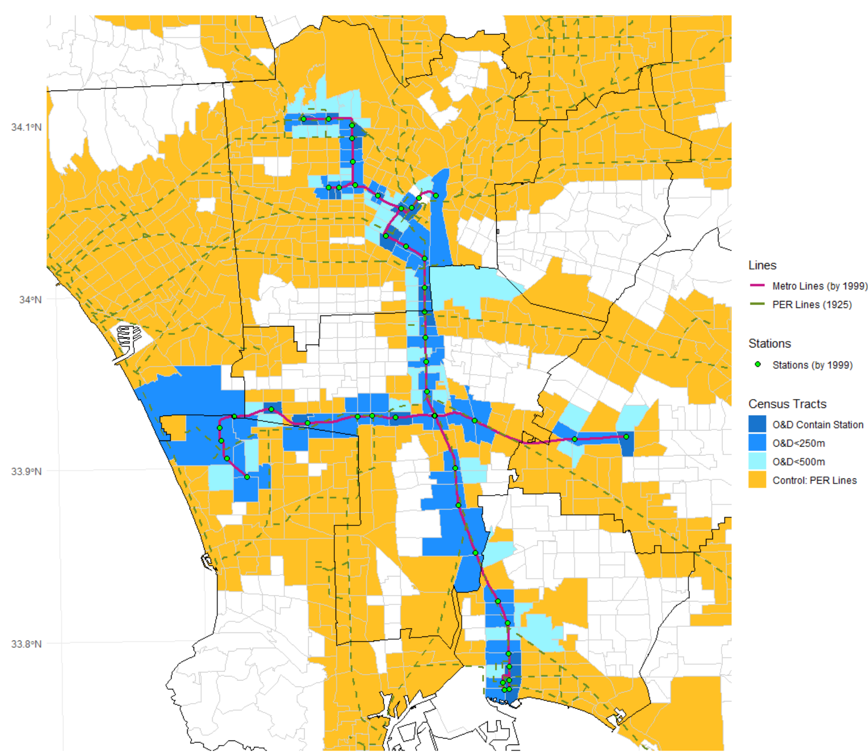


Figure 3: Comparison of Commuting and Market Access measures

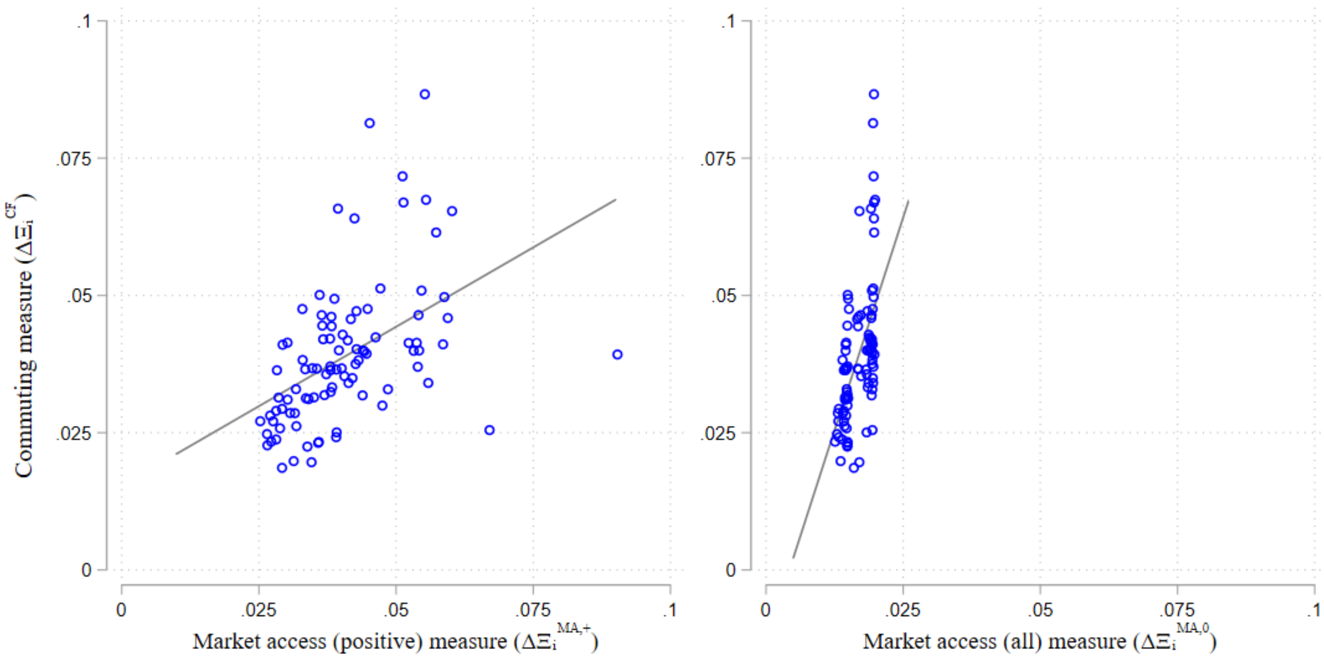


Figure 4: Does $\omega = \epsilon w$, and if not, what is it capturing?

Figure 5: Take-up of LA Metro Rail for commuting does not vary by income, but overall take-up of transit (including bus) does.

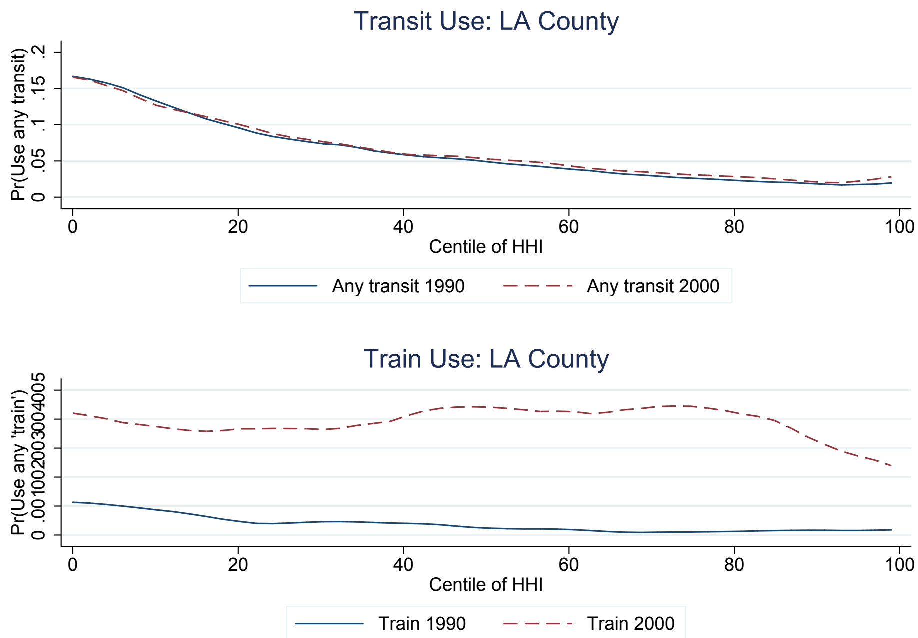


Table 1: Descriptive statistics on transportation in Los Angeles and station placement

	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 only 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

	(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: Effect of Transit on Commuting Flows (by 2000) - Adjacencies

	(1)	(2)	(3)
O & D contain station	0.087*	0.171**	0.170**
	(0.041)	(0.060)	(0.060)
O & D <250m from station	-0.022	-0.032	-0.032
	(0.046)	(0.059)	(0.059)
O & D <500m from station	-0.015	-0.049	-0.049
	(0.039)	(0.050)	(0.050)
O & D <1000m from station			-0.005
			(0.031)
Adjacency Base	500m	1000m	1000m
Tract Pair FE	Y	Y	Y
Group-X-YrFE	Y	Y	Y
N	217517	301662	301662

High-dimensional fixed effects estimates of λ^D estimated from adjacencies (see text). Treatment variables are mutually exclusive with others in each column. All estimates include tract pair and group-by-year 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: 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 5: Comparing commuting flow and market access measures

	κ	Mean	S.D.	Min	Max
Ξ_i^{CF}	-	.0390	.0132	.0186	.0866
$\Xi_i^{MA,+}$	-0.05	.0408	.0108	.0253	.0904
$\Xi_i^{MA,+}$	-0.02	.0379	.0099	.0232	.0818
$\Xi_i^{MA,+}$	-0.01	.0368	.0096	.0220	.0788
$\Xi_i^{MA,+}$	-0.005	.0362	.0094	.0213	.0773
$\Xi_i^{MA,0}$	-0.05	.0169	.0024	.0126	.0199
$\Xi_i^{MA,0}$	-0.02	.0133	.0018	.0107	.0152
$\Xi_i^{MA,0}$	-0.01	.0121	.0016	.0099	.0136
$\Xi_i^{MA,0}$	-0.005	.0114	.0015	.0095	.0128

There are 93 treated observations in each case, simulated using the statistically significant estimates Table 2, Column 5, in the Subway Plan (All) Sample.

Table 6: IV estimates of labor supply elasticity (ϵ)

	$\hat{\omega}_{jt}$ (1)	$\hat{\omega}_{jt}$ (2)	$\hat{\omega}_{jt}$ (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 ($\hat{\omega}_j t$) estimated from a panel specification in columns 1 and 3, relying on ij fixed effects to control for distance. Column 2 uses $\hat{\omega}_j t$ estimated year-by-year, using network distance to control for distance. Standard errors clustered by tract in parentheses: $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 7: IV estimates of inverse housing supply elasticity (ψ)

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

	(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: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 9: IV estimates of inverse labor demand elasticity ($\alpha - 1$)

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

Table 10: Transit and non-commuting fundamentals (other effects of transit)

	(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 thirty-two 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 11: 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%, 30yy)		-$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 12: Dynamic effects of transit on flows (2002-15), Linear

	(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 12 – continued from previous page

	(1)	(2)	(3)	(4)	(5)	(6)
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

December 2018

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) Origen-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 and Company (1925). I georeference this map in ArcGIS, and then processe 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#. See Figure **??** to reference data processing.

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 this 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 accords reasonably well with ad hoc probes of quality. However, there are instances of extreme values that becomes 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 and 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 lies 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 and 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, (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 item unobservable to route planning GIS systems that may induce measurement error.
4. CTPP location data is 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 does 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

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: Rearranging the system in Equations (6), (9), (12), (13), and (11) into a more convenient form gives:

$$W_j^{\frac{1+\epsilon(1-\alpha)}{1-\alpha}}\Omega_j = \bar{N}^{-1}K_{0j}\sum_s W_s^{\epsilon}\Omega_s$$
$$\Omega_j = \sum_r K_{1rj}Q_r^{-\epsilon(1-\zeta)}$$
$$Q_i^{-\epsilon(1-\zeta)-\frac{1+\psi}{\psi}}\Phi_i = \bar{N}^{-1}K_{2i}\sum_s W_s^{\epsilon}\Omega_s$$
$$\Phi_i = \sum_s K_{1is}W_s^{\epsilon+1}$$

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 (14).[B.1]

Proposition 2

Existence: 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 (14) 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]

Proposition 3

Under the assumption that $e^{-\kappa \delta_{ij}}$ are fixed terms, $\Delta z_{it}^{HD,X}$, $\Delta z_{(-i)jt'}^{HS,X}$, and $\Delta z_{it}^{LS,X}$ are linear combinations of $\Delta z_{jt}^{LD,X}$. Therefore, A2 maps directly into M2, and A4 maps directly into M4. Condition M3 imples $\mathbb{E}[\Delta z_{j't}^{LD,X} \Delta \ln(B_{it} D_{ijt})] = 0, \; \forall \, i, j' \neq j$, which holds given M1.

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

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

First, I show that the expression in Equation (25) has an equivalent log-sum representation. Begin by dividing counterfactual and factual expected utilities (from Equation 7):

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

where $\{ij\} = \{rs\}$ track summation sets. Substituting in Equation (6) for some particular ij into the above twice (once for π'_{ij} and once for π_{ij}) and taking logs gives Equation (25).

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

$$\mathbb{E}[\bar{\mathcal{W}}'] - \mathbb{E}[\bar{\mathcal{W}}] = \frac{1}{\mu} \ln\left(\frac{\sum_{k \in K_1} e^{V'_k}}{\sum_{k \in K_0} e^{V_k}}\right) \tag{F-2}$$

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

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

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 (25) can be used for welfare

B.4. Thanks to Wei You for noting that (25) and a log-sum expression are interchangeable:

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

evaluation when $\epsilon \in (0, 1]$ and well as $\epsilon > 1$.

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). Spreadsheet available by request.

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

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

D Counterfactual Estimation

First, note that the following hold:

$$\hat{W}_i = \hat{A}_i \hat{\bar{N}}^{\alpha-1} \left(\frac{\sum_r \pi_{ri}\hat{\pi}_{ri}}{\sum_r \pi_{ri}} \right)^{\alpha-1}$$

$$\hat{Q}_i = \hat{C}_i^{1/(1+\psi)} \left(\frac{\hat{\bar{N}} \sum_s \pi_{is}\hat{\pi}_{is} W_s \hat{W}_s}{\sum_s \pi_{is} W_s} \right)^{\psi/(1+\psi)}$$

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

where $\hat{\bar{N}} = 1$ 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:

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

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 $\hat{\bar{N}}^{(0)}$ 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 $\hat{\hat{N}}^{(t-1)}$, $\{\hat{W}_i^{(t-1)}\}$, $\{W_i\}$, $\{\hat{\pi}_{ij}^{(t-1)}\}$, and $\{\pi_{ij}\}$

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

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

   (d) Define $\hat{\hat{N}}^{(temp)}$ 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 $\hat{X} \in \{\hat{Q}, \hat{W}, \hat{\pi}, \hat{\hat{N}}\}$, where ξ is a weight that disciplines updating.

   (f) Estimate movement as:

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

   (g) Stop when movement is below convergence criterion

E Identification under alternative assumptions

In this section, I discuss identification under more general settings that those in the simplest version of the model presented in the paper. The two modification I consider are: (i) endogenous land use determination (no zoning), and (ii) the presence of agglomeration and other forces. As discussed in the paper, it is unlikely that either of these plays a significant role in the environment presented in the paper. Nonetheless, it is illustrative to work through these variations.

I first display the identification assumptions from the main text. Though these will continue to be necessary, they will not be sufficient.

$$\mathbb{E}[\Delta z_{jt}^{LD,R} \times \Delta \ln(E_{jt} D_{ijt})] = 0, \ \forall \, ij \tag{A-1}$$

$$\mathbb{E}[\Delta z_{jt}^{LD,R} \times \Delta \ln(C_{it})] = 0, \ \forall \, ij \tag{A-2}$$

$$\mathbb{E}[\Delta z_{j't}^{LD,R} \times \Delta \ln(B_{it} D_{ijt})] = 0, \ \forall \, ij' \neq ij \tag{A-3}$$

$$\mathbb{E}[\Delta z_{j't}^{LD,R} \times \Delta \ln(A_{jt})] = 0, \ \forall \, j' \neq j \tag{A-4}$$

and their simplified versions that accomodate the presence of rich fixed effects:

$$\mathbb{E}[\Delta z_{jt}^{LD,R} \times \Delta \ln(E_{jt})] = 0, \ \forall \, j \tag{A-1a}$$

$$\mathbb{E}[\Delta z_{jt}^{LD,R} \times \Delta \ln(C_{it})] = 0, \ \forall \, i \neq j \tag{A-2a}$$

$$\mathbb{E}[\Delta z_{j't}^{LD,R} \times \Delta \ln(B_{it})] = 0, \ \forall \, i \tag{A-3a}$$

Below, I describe additional conditions for identification, and their plausibility, under various changes to model form and data availability.

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):

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

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

where k here represent distance kernals and $N_i^H = \bar{N}\sum_s \pi_{is}$ 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, according to the estimates in Table **??**. This is higher that 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.1] 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.2]

E.3 Agglomeration and Endogenous Land Use

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

E.1. Note that this implies $\Delta L_{it}^Y = -\Delta L_{it}^H$.

E.2. 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

F.1 Commuting calculations

Gravity estimates

First, a note on measure of travel time/cost:

- Network Travel Cost: calculated from a road network using GIS software. This only available in the cross-section, but is available for all pairs.
- Travel Time: average pair travel time from the CTPP. Panel data, but only available between pairs with positive commuting that satisfy disclosure requirements.

The Network Travel Cost (or similar variants using Google Maps) is mostly standard in the new EG literature. Table F14 shows that to translate between the two, roughly divide Network Travel Cost by 2.

Table F11 provides cross-sectional estimates of travel time disutility using several estimators and both years. Estimates are roughly consistent with the previous literature, and suggest that the Travel Time–Network Travel Cost conversion is reasonable.

Table F12 compares estimates of:

$$n_{ijt} = \omega_{jt} + \theta_{it} - \epsilon\kappa\tau_{ijt} + \ln(D_{ijt})$$
$$n_{ijt} = \omega_{jt} + \theta_{it} - \epsilon\kappa\tau_{ijt} + \varsigma_{ij}^{D} + \ln(D_{ijt})$$

The second of these is particularly prone to measurement error in reported or modeled travel times, because the fixed effect means that any autocorrelation in error magnifies the issue.

Table F13 first estimates ς_{ij}^{D}, then runs:

$$\varsigma_{ij}^{D} = -\epsilon\kappa\tau_{ij} + u_{ij}$$

for a particular t.

Permanence of unobserved determinants of commuting flows

Roughly half of the variation in commuting flows is due to unobserved, time-invariant characteristics. This details the analysis that leads to this conclusion. First, I run gravity models of commuting in the cross section with and without travel time:

$$n_{ij} = \omega_{j} + \theta_{i} - \epsilon\kappa\tau_{ij} + \ln(D_{ij})$$

The inclusion of travel time increases the R^2 from 0.20 to 0.26 in 1990, and 0.17 to 0.21 in 2000. This suggests that while travel time plays an important role, other factors are important.

Second, I run the panel gravity model with and without pair fixed effects:

$$n_{ijt} = \omega_{jt} + \theta_{it} + \varsigma_{ij}^{D} + \ln(D_{ij})$$

The model without pair fixed effects has an R^2 of 0.18 (or 0.21 if restricted to observations that have two non-zero commuting flows), but jumps to 0.81 with pair fixed effects. This regression excludes travel time. Regressing the pair fixed effects on travel time or network-based measures of travel cost:

$$\widehat{\varsigma_{ij}^{D}} = \epsilon\kappa\tau_{ij} + \ln(D_{ij})$$

lead to R^2 values of between 0.07 and 0.19.

Putting this together, pair fixed effects explain about 60% of the variation in the panel, conditioning on origin- and destination-by-year fixed effects. Only about 20% of these fixed effects are explained by commuting time, meaning 80% are not. Putting these together, we see that about 48% of commuting flows are explained by time-invariant, non-distance characteristics.

Appendix Figures and Tables

Figure F1: Ridership, 1990-2000

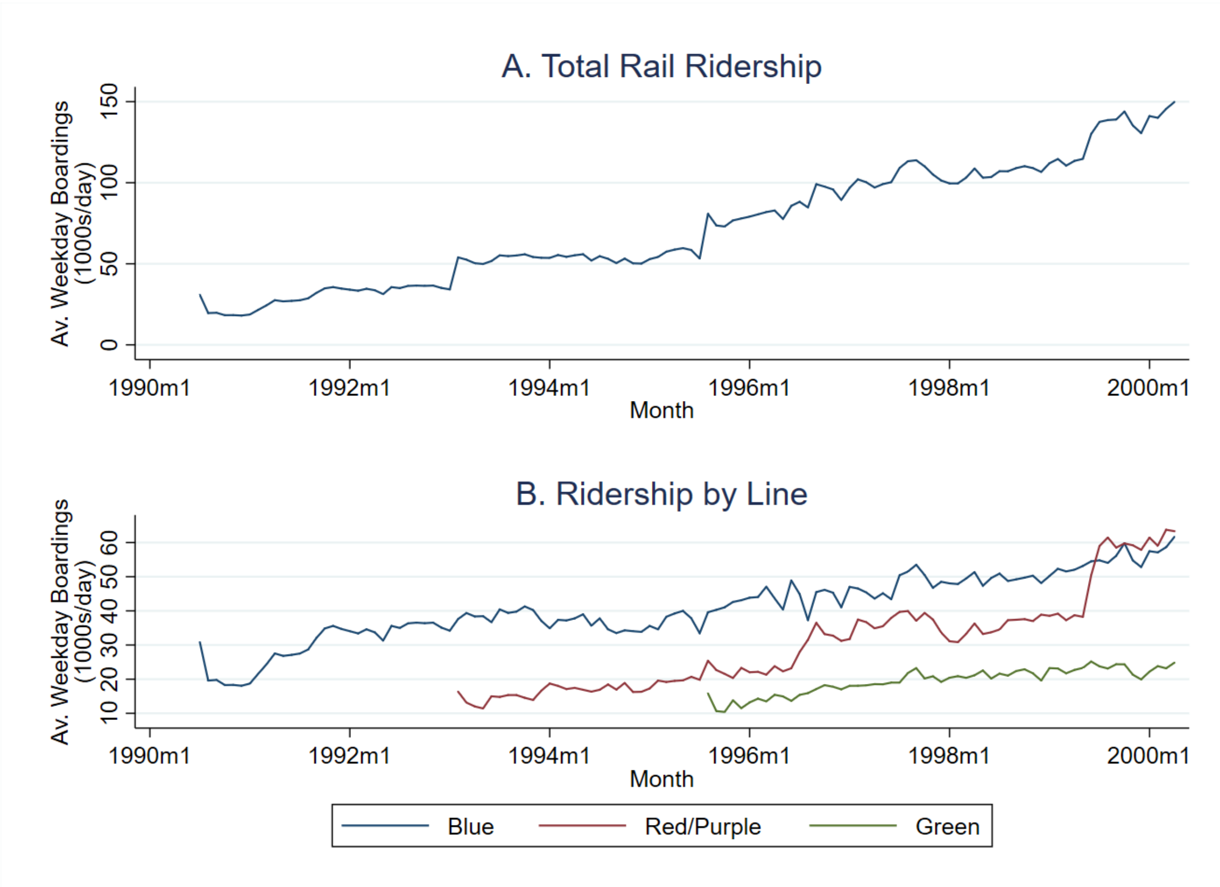


Figure F2: Ridership, 1990-2014

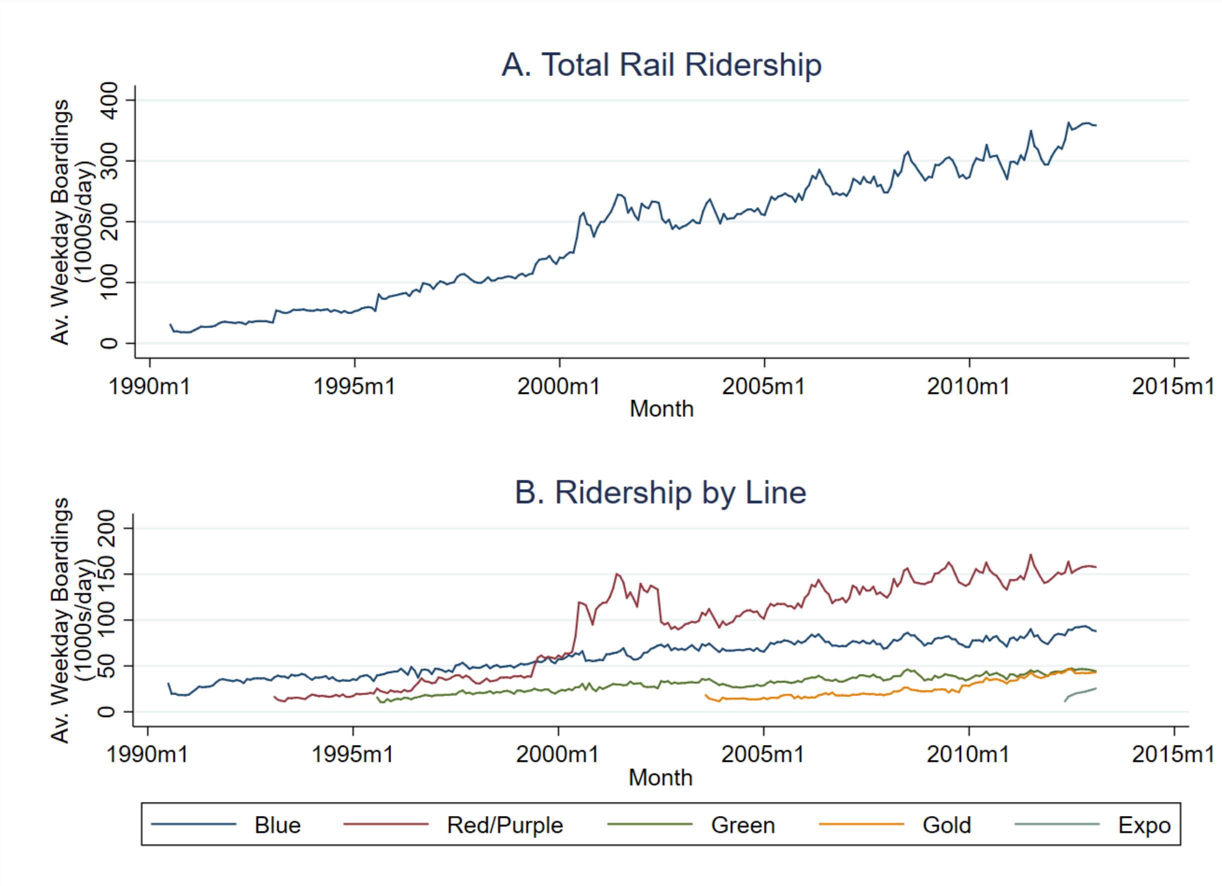


Figure F3: Glossary of variables and parameters

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 and Co. develop 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 tract-level commuting behavior, 1970-1990

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Commuting by automobile								
$\text{Proximity}_i^{500m} \times t$	-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. Transit (rail and bus) commuters, >0								
$\text{Proximity}_i^{500m} \times t$	-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
C. Transit (rail and bus) commuters, all								
$\text{Proximity}_i^{500m} \times t$	-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
D. No car households								
$\text{Proximity}_i^{500m} \times t$	-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
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 thirty-two total. Estimates show pre-trends from 1970-1990 for tracts treated by 1999, except for Panel D, which only covers 1980-1990. Panel A, C, D estimated by PPML with exposure set to relevant tract population. Panel B uses log commuters in a linear specification. All regressions include tract fixed effects. Standard errors clustered by tract in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F2: 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 F3: Differences in initial levels

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Linear								
O & D <500m from station	0.213**	0.193**	0.148**	0.134**	0.132**	0.127**	0.148**	0.150**
	(0.040)	(0.042)	(0.037)	(0.033)	(0.040)	(0.035)	(0.038)	(0.033)
N	9619	9611	37023	37020	49537	49527	145440	145239
PPML								
O & D <500m from station	0.328**	0.296**	0.256**	0.227**	0.223**	0.206**	0.224**	0.243**
	(0.044)	(0.046)	(0.045)	(0.038)	(0.045)	(0.039)	(0.046)	(0.038)
N	34809	34804	154852	154852	203249	203232	628704	627824
Control Network	1925 Imm	1925 Imm	1925 All	1925 All	PER Lines	PER Lines	All	All
POW FE	Y	Y	Y	Y	Y	Y	Y	Y
RES FE	Y	Y	Y	Y	Y	Y	Y	Y
Sbcty-X-Sbcty FE	-	Y	-	Y	-	Y	-	Y

High-dimensional fixed effects estimates of ex-ante (1990) differences between tracts that become treated and control tracts. All control networks are 'loose' (see text). All estimates include tract of work and tract of residence fixed effects. Standard errors clustered by FIND OUT: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F4: Do higher initial levels lead to higher growth

	Linear				PPML			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$N_{ij,1990}$	-0.014**	-0.018**			-0.004**	-0.004**		
	(0.001)	(0.002)			(0.001)	(0.001)		
$\ln(N_{ij,1990})$			-0.473**	-0.593**			0.065**	-0.196**
			(0.008)	(0.006)			(0.012)	(0.012)
N	291000	290580	291000	290580	1259500	1256986	764988	758066
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
Sbcty-X-Sbcty-X-Yr FE	-	Y	-	Y	-	Y	-	Y
Highway Control	-	Y	-	Y	-	Y	-	Y

High-dimensional fixed effects estimates of lagged level or log-level of flow. Control network is '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 F5: 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 F6: 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 F7: 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. 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 F8: Effect of Transit on Commuting Flows (by 2000) - Adjacencies, PPML

	(1)	(2)	(3)
O & D contain station	0.029	0.113	0.113
	(0.066)	(0.094)	(0.094)
O & D <250m from station	-0.020	-0.003	-0.003
	(0.045)	(0.060)	(0.060)
O & D <500m from station	-0.015	-0.010	-0.010
	(0.039)	(0.048)	(0.048)
O & D <1000m from station			0.038
			(0.033)
Adjacency Base	500m	1000m	1000m
Tract Pair FE	Y	Y	Y
Group-X-YrFE	Y	Y	Y
N	341794	484533	484533

High-dimensional fixed effects estimates of λ^D estimated from adjacencies (see text). Treatment variables are mutually exclusive with others in each column. All estimates include tract pair and group-by-year 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 F9: IV estimates of labor supply elasticity (ϵ), untrimmed

	$\hat{\omega}_{jt}$ (1)	$\hat{\omega}_{jt}$ (2)	$\hat{\omega}_{jt}$ (3)
$\ln(W_{jt})$	1.015^{+}	2.876^{**}	2.968^{**}
	(0.551)	(0.844)	(0.838)
F-stat (KP)	13.591	17.430	17.430
$\hat{\omega}$ estimated:	Linear, Panel	PPML Yr-by-yr	PPML Panel
N	2427	2528	2528

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 not trimmed. Column 1 uses a linear specification, columns 2 and 3 assume a Poisson model. Place of work-by-year fixed effects ($\hat{\omega}_j t$) estimated from a panel specification in columns 1 and 3, relying on ij fixed effects to control for distance. Column 2 uses $\hat{\omega}_j t$ estimated year-by-year, using network distance to control for distance. Standard errors clustered by tract in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F10: Transit and non-commuting fundamentals (other effects of transit), robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Effect on productivity ΔA, $\alpha - 1 = -0.25$								
$\text{Proximity}_i^{500m} \times t$	-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, $\epsilon(1-\zeta) = 0.125$								
$\text{Proximity}_i^{500m} \times t$	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, $\psi = 2.292$								
$\text{Proximity}_i^{500m} \times t$	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, $\epsilon = 1$								
$\text{Proximity}_i^{500m} \times t$	-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, $\epsilon = 0.498$								
$\text{Proximity}_i^{500m} \times t$	-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 forty 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 F11: Estimating travel costs from single year cross-sections

	1990				2000			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Travel Time	-0.0108**	-0.0244**			-0.0085**	-0.0200**		
	(0.0003)	(0.0008)			(0.0002)	(0.0007)		
Network Travel Cost			-0.0187**	-0.0528**			-0.0186**	-0.0510**
			(0.0004)	(0.0019)			(0.0003)	(0.0010)
N	291000	290580	291000	290580	1259500	1256986	764988	758066
Model	Linear	PPML	Linear	PPML	Linear	PPML	Linear	PPML
POW FE	Y	Y	Y	Y	Y	Y	Y	Y
RES FE	Y	Y	Y	Y	Y	Y	Y	Y

Adjust: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F12: Estimating travel costs from the panel

	(1)	(2)	(3)	(4)	(5)	(6)
Travel Time	-0.0095**	-0.0003*		-0.0220**	0.0001	
	(0.0002)	(0.0001)		(0.0007)	(0.0002)	
Network Travel Cost			-0.0186**			-0.0518**
			(0.0003)			(0.0013)
N	775721	291000	774602	785883	311334	1257775
Model	Linear	Linear	Linear	PPML	PPML	PPML
Tract Pair FE	-	Y	-	-	Y	-
POW-X-Yr FE	Y	Y	Y	Y	Y	Y
RES-X-Yr FE	Y	Y	Y	Y	Y	Y

Adjust: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F13: Estimating travel costs from a two step estimation

	(1)	(2)	(3)	(4)	(5)	(6)
Travel Time	-0.0143**	-0.0109**		-0.0179**	-0.0149**	
	(0.0004)	(0.0003)		(0.0004)	(0.0004)	
Network Travel Cost			-0.0198**			-0.0244**
			(0.0008)			(0.0005)
N	145500	145500	145440	382697	402811	628712
Model	Linear	Linear	Linear	PPML	PPML	PPML
Year	1990	2000	-	1990	2000	-

Standard errors clustered tract of residence and tract of work in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F14: Converting between Network Travel Cost and Travel Time

	Travel Time (1)	Travel Time (2)	Travel Time (3)	Travel Time (4)
Network Travel Cost	0.4967**	0.5394**	0.4408**	0.4887**
	(0.0026)	(0.0103)	(0.0021)	(0.0086)
N	382281	382244	402524	402516

EXPLAIN $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F15: Testing parallel pre-trends in other (non-commuting) tract characteristics, 1970-1990

	ln Res. Emp. (1)	ln #HHs (2)	ln HHI (3)	ln House Value (4)	% Coll. Grads (5)	Pov. Rate (6)	% Moved <5yrs (7)
Subway Plan (Immediate) Sample							
$\text{Proximity}_i^{500m} \times t$	0.029	-0.011	-0.013	-0.002	-0.008*	0.008+	-0.011*
	(0.020)	(0.017)	(0.013)	(0.019)	(0.004)	(0.005)	(0.006)
N	1629	1629	1628	1555	1629	1629	1629
Subway Plan (All) Sample							
$\text{Proximity}_i^{500m} \times t$	0.012	-0.031+	-0.019	-0.017	-0.013**	0.012**	-0.014**
	(0.020)	(0.017)	(0.012)	(0.018)	(0.003)	(0.004)	(0.005)
N	3786	3786	3779	3688	3786	3786	3786
PER Sample							
$\text{Proximity}_i^{500m} \times t$	0.002	-0.032+	-0.020	-0.034+	-0.015**	0.013**	-0.014**
	(0.021)	(0.017)	(0.013)	(0.018)	(0.004)	(0.004)	(0.005)
N	4631	4629	4619	4502	4631	4632	4631
Full Sample							
$\text{Proximity}_i^{500m} \times t$	0.025	-0.027	-0.016	-0.022	-0.015**	0.015**	-0.016**
	(0.020)	(0.017)	(0.012)	(0.017)	(0.003)	(0.004)	(0.005)
N	11651	11641	11567	11407	11657	11733	11659
Tract FE	Y	Y	Y	Y	Y	Y	Y
Sbcty-X-Yr FE	Y	Y	Y	Y	Y	Y	Y

Each column of each panel presents the results of a different regression, for twenty-eight 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 F16: Control group validity: labor demand shocks and treatment status

			1925 Plan Sample				PER Lines Sample			
	$z_{jt}^{LD,e}$ (1)	$z_{jt}^{LD,w}$ (2)	$z_{jt}^{LD,e}$ (3)	$z_{jt}^{LD,e}$ (4)	$z_{jt}^{LD,w}$ (5)	$z_{jt}^{LD,w}$ (6)	$z_{jt}^{LD,e}$ (7)	$z_{jt}^{LD,e}$ (8)	$z_{jt}^{LD,w}$ (9)	$z_{jt}^{LD,w}$ (10)
A. Tract centroid within 500 meters of station										
1[Transit]	-0.004	-0.002	-0.006	0.000	-0.003	0.000	-0.006	0.002	-0.002	0.001
	(0.006)	(0.003)	(0.006)	(0.007)	(0.003)	(0.003)	(0.006)	(0.007)	(0.003)	(0.003)
N	5,074	5,074	1,422	1,422	1,422	1,422	1,884	1,880	1,884	1,880
B. Any part of tract within 500 meters of station										
1[Transit]	-0.002	-0.003	-0.004	0.005	-0.004*	-0.000	-0.005	0.005	-0.004*	0.000
	(0.004)	(0.002)	(0.004)	(0.005)	(0.002)	(0.002)	(0.004)	(0.005)	(0.002)	(0.002)
N	5,074	5,074	1,458	1,458	1,458	1,458	1,924	1,920	1,924	1,920
Tract FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Subcty-x-yr FE	-	-	-	Y	-	Y	-	Y	-	Y

Each column of each panel presents the results of a different regression of the labor demand shock (measured in wage or employment) on treatment status, for twenty total. Regressions include year and tract fixed effects, and some include subcounty-by-year fixed effects. Standard errors clustered by tract in parentheses: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table F17: 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 sixteen 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: $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$