6-3-2020

Where do the poor live in cities? Revisiting the role of public transportation on income sorting in US urban areas

Erik Nelson
Bowdoin College, enelson2@bowdoin.edu

Where do the poor live in cities? Revisiting the role of public transportation on income sorting in US urban areas

Erik Nelson[1]

Abstract: Glaeser et al. (2008) argue that the relative distribution of poor and rich households (HHs) in American cities is "strongly" explained by the spatial location of the cities' public transportation (PT) networks. Among their claims: 1) The broad distribution of poor and rich HHs in the typical American city is consistent with a basic monocentric city model that includes commute technology speeds; 2) Poor commuters will overwhelmingly transition from commuting by PT to car if they experience a substantial increase in their HH's income; 3) areas in American cities that receive new PT infrastructure become poorer over time. Using 2017 data I find empirical evidence that partially or wholly contradicts these three claims. First, as of 2017, the observed concentration of poor HHs in the inner city and rich HHs in the suburbs of the US' smaller cities cannot be explained by monocentric model that includes commute speeds. Second, as of 2017, significant increases in poor HHs' incomes were not expected to lead to a "massive shift" towards car commuting in these HHs; most of these poor workers commute by car already. Third, using data from four cities that expanded their light-rail and rapid-bus network in the early 2000s, I find that neighborhoods surrounding new light-rail or rapid-bus stations either saw little change in their income patterns or became slightly richer after station opening. In conclusion, as of 2017, the spatial distribution of HH incomes within American urban areas is not as intricately linked to the location of PT networks as Glaeser et al. (2008) would have us believe. As an addendum to the analysis I add some thoughts on how the COVID-19 pandemic might affect commuting behavior and income distributions within urban areas over the next decade.

Keywords: Monocentric city model; public transportation; commuting; household income distribution; mode choice; random utility model; event-study; light rail

JEL codes: O18; R2; R41

[1] Department of Economics, Bowdoin College, Brunswick, ME

1. Introduction

The relationship between household (HH) income and distance from the central business district (CBD) in a US urban area can generally be described by one of two patterns. In the first, HH income generally increases in distance from the CBD. The Los Angeles (LA) core based statistical area (CBSA) typifies this type of urban HH income gradient (Fig. 1). In the other general pattern, a ring of poor HHs lie between two rings of richer HHs where one ring of more well-to-do HHs is centered on the CBD and another lies in the urban area's farther-out suburbs. Chicago typifies this type of urban household income gradient (Fig. 1).

1a. A monocentric city model without and with commuting technology

Leroy and Sonstelie (1983) and Glaser et al. (2008) have described a monocentric city model that can produce HH income sorting consistent with LA's or Chicago's general patterns (or at least the first two rings in a Chicago-like city). Whether the model produces a LA or Chicago-like city largely depends on 1) what commute technologies poorer commuters tend to use versus richer commuters and 2) the speeds of the different commute technologies.

Assume a HH with one worker is deciding how far to live from an urban area's CBD. The HH has two major concerns in its location choice: the cost of their lot and their worker's cost of commuting to their job in the CBD (Leroy and Sonstelie 1983) (I assume amenities, another potential factor in location choice, are spread evenly across the urban area and therefore not a factor in location choice.) Let $P(D)$ indicate the daily price of a square foot of a lot that is located D miles from the CBD and let A indicate the square footage of each lot in the urban area.

Further, let W be the worker's opportunity cost of a minute of commuting time and T indicate the minutes needed by the worker to commute a mile.

The utility-maximizing HH will to choose to live at the D that minimizes the sum of its daily costs (Glaeser et al. 2008),

$$\min_D C = AP(D) + 2(WTD) \quad (1.1)$$

where $AP(D)$ is the HH's daily expenditure on a lot of size A at D and $2(WTD)$ is the HH's daily cost of the commuting to and from work given their choice of D. The HH will minimize its daily cost by choosing to live at the D that satisfies,

$$P'(D) = -2WT/A \quad (1.2)$$

The HH's location choice indifference contour or bid-rent curve in P-D space has a slope of negative $2WT/A$ and indicates every P and D combination that generates the HH's minimized C. Assuming this urban area contains identical HHs and only has one commute technology (i.e., each household has the same T), this area will reach location equilibrium when each HH selects a lot at arbitrary D and pays $P(D)$ per square foot according to the city's one bid-rent curve.

Now assume an urban area with both low-income (or poor) HHs and high-income (or rich) HHs where all HHs within an income class are identical. Let Y_{Poor} and Y_{Rich} indicate the incomes at poor and rich HHs, respectively. Assume the rich worker's opportunity cost of commuting is greater due to their higher hourly incomes (i.e., $W_{Rich} > W_{Poor}$) and landlords design lots for the rich HHs that are different than those designed for the poor HHs (i.e., $A_{Rich} \neq A_{Poor}$; Glaeser et al 2008). However, I still assume T is identical for both types of HHs. If I assume the rich HH bid-rent curve is steeper than the poor HH curve (Fig. 2) then, in equilibrium, rich

HHs will out-bid poor HHs for a lot closer to the CBD, landlords will place lots of size A_{Rich} near the CBD and lots of size A_{Poor} farther from the CBD, and the following will hold,

$$\underbrace{\frac{\Delta W}{\Delta I} \times \frac{Y_{Poor}}{W_{Poor}}}_{\varepsilon_Y^W} > \underbrace{\frac{\Delta A}{\Delta I} \times \frac{Y_{Poor}}{A_{Poor}}}_{\varepsilon_Y^A} \tag{1.3}$$

where $\Delta W = W_{Rich} - W_{Poor}$, $\Delta A = A_{Rich} - A_{Poor}$, and $\Delta I = Y_{Rich} - Y_{Poor}$ (see SI Text 1 for an explanation of why (1.3) must hold if the rich HHs concentrate in inner ring of the urban area). The left-hand side of (1.3) is the urban area's elasticity of the time cost of commuting with respect to income (i.e., for every 1% increase in income, the cost of commuting increases by ε_Y^W percent) and the right-hand side of (1.3) is the urban area's income elasticity of demand for lot size (i.e., for every 1% increase in income, the demand for larger lots increases by ε_Y^A percent).

Empirical research suggests that inequality (1.3), as well as my assumptions $W_{Rich} > W_{Poor}$ and $A_{Rich} \neq A_{Poor}$, held in the representative US urban area: as of the early 2000s, ε_Y^W was estimated to be approximately 0.75 and ε_Y^A was estimated to be in the range from 0.25 to 0.50 (Glaeser et al. 2008). Therefore, *as presented so far*, a basic monocentric city model with commute technology parametrized with representative urban elasticities *cannot* explain a city like LA where the poor concentrate around the core (the first band or ring around the CBD is relatively poor) and the rich in the suburbs (the second band or ring around the CBD is relatively richer).

However, if the poor commuters in the modeled urban area use a sufficiently slower transportation technology than the area's rich commuters then the modeled urban can mimic LA's broad pattern of HH income sorting. In other words, unlike the monocentric city model where all commuters travel the same speed, the version of the model where rich and poor commuters can travel at different speeds supports two sorting equilibria (Fig. 3). Assuming $\varepsilon_Y^W \approx 0.75$ and $\varepsilon_Y^A =$[0.25 to 0.50], if the income gap between the rich and poor HHs in this urban area (i.e., $Y_{Rich} - Y_{Poor}$) is not too large and the poor HH's commute technology is sufficiently slower than rich HH's commute technology (i.e., $T_{Poor} \gg T_{Rich}$) then $\frac{W_{Rich}T_{Rich}}{W_{Poor}T_{Poor}} < \frac{A_{Rich}}{A_{Poor}}$ or,

$$\varepsilon_Y^A + \frac{T_{Poor}-T_{Rich}}{T_{Poor}}\left(\frac{Y_{Poor}}{Y_{Rich}-Y_{Poor}} + \varepsilon_Y^W\right) > \varepsilon_Y^W \tag{1.4}$$

holds and the poor HHs concentrate in the first ring (see SI Text 1). Otherwise, if $Y_{Rich} - Y_{Poor}$ is large, $T_{Poor} \approx T_{Rich}$, or there is only commute technology in the urban then (1.4) does not hold and the rich concentrate in the first ring.

A common criticism of this model is its lack of other spatial phenomenon that affect the spatial pattern of HH incomes in an urban area such as crime, pollution, and schooling quality and discriminatory tastes. However, Glaeser et al. (2008) argue that these factors largely explain the *separation* of the poor and the rich in an urban area but not the spatial order of poor and rich enclaves within an urban area. "A satisfying theory of urban centralization should explain not only why the poor and the non-poor live apart, but also why, conditional upon the poor and non-poor living apart, the poor choose to live closer to the city center." (p. 7). Here I assume the poor and rich generally live apart. I am trying to explain why in some cities poor HHs tend to cluster in the first ring of the city and in others they cluster in a second ring.

1b. Assigning commuting technologies to income classes

Assume there are two transportation technologies in our modeled urban area, the car and public transportation (PT). I can assume parameter T_{Rich} is equal to T_{Car} if rich HHs find commuting with a car their least costly option given their lot location choice *D*,

$$\underbrace{C + W_{Rich}T_{Car}D}_{\substack{\text{Cost of a rich HH} \\ \text{commute in a car}}} < \underbrace{Q + W_{Rich}T_{PT}D + W_{Rich}F}_{\substack{\text{Cost of a rich HH} \\ \text{commute on PT}}} \tag{1.5}$$

where *C* is the fixed cost of a car per commute, *Q* is the fare for a commute by PT, and *F* is the PT wait and egress time (in minutes) per commute (relative to car commuting). According to (1.5), if $C < P + W_{Rich}F$ then a rich commuter with a lot at $D = 0$ will use a car. If, in addition, $T_{PT} > T_{Car}$ then (1.5) will hold at all $D > 0$ because $W_{Rich}T_{PT}D > W_{Rich}T_{Car}D$ at all *D*.

Further, I can convert T_{Poor} to T_{PT} if the poor commuter finds PT their least costly option given their lot location choice *D.* Assume $D = D^*$, the furthest extent of the city. A poor HH commuter living at this point will use PT if,

$$\underbrace{C + W_{Poor}T_{Car}D^*}_{\substack{\text{Cost of a poor HH} \\ \text{commute in a car}}} > \underbrace{P + W_{Poor}T_{PT}D^* + W_{Poor}F}_{\substack{\text{Cost of a poor HH} \\ \text{commute on PT}}} \tag{1.6}$$

or

$$C > P + D^*(T_{PT} - T_{Car})W_{Poor} + W_{Poor}F \tag{1.7}$$

(assuming PT is available at D^*). Further, note that if (1.7) holds at D^*then then it will hold at $D < D^*$, including $D = 0$.

1c. Glaeser et al. (2008)'s analysis of income spatial patterns in US cities

Using data from the early 2000s, Glaeser et al. (2008) find the following model parameter values for the *average* American urban area: $T_{PT} = 3$ minutes mile^{-1}, $T_{Car} = 1.6$ minutes mile^{-1}, $\varepsilon_Y^A = [0.25, 0.50]$, and $\varepsilon_Y^W \approx 0.75$. Further, by analyzing plots of average census tract (CT)-level incomes by distance from the CBD for New York, Philadelphia, Chicago, Atlanta, Phoenix, and Los Angeles they find that $\frac{Y_{Poor}}{Y_{Rich}-Y_{Poor}}$ was never less than 0.7 in the early 2000s. Finally, the authors also use a back-of-the-envelope analysis to find that the representative poor commuter has incentive to use PT and the representative rich commuter has incentive to use a car in the typical US urban area (SI Text 2). (Their analysis suggests a poor worker that transitions from making \$10 or less per hour (2001 USD) to \$15 to \$20 per hour (2001 USD) will transition from PT to car commuting.)

Glaeser et al. (2008) then show that these estimated parameters support the inequality,

$$\varepsilon_Y^A + \frac{T_{PT}-T_{Car}}{T_{PT}}\left(\frac{Y_{Poor}}{Y_{Rich}-Y_{Poor}} + \varepsilon_Y^W\right) > \varepsilon_Y^W \tag{1.8}$$

where $T_{Poor} = T_{PT}$ and $T_{Rich} = T_{Car}$. In other words, according to the basic monocentric city model with different commuting technologies, the typical urban US area should have an inner ring of generally poorer commuters using PT and a second ring of generally richer commuters using cars (e.g., Los Angeles). Their data and analysis also suggest that the tendency of the poor to concentrate in the inner ring will be amplified if the urban area's PT network is not dense in the second ring.[1]

Glaeser et al. (2008) then proceed to show that observed patterns of HH income sorting in US urban areas are consistent with their theoretical findings. First, using regression analysis, they find that US census tracts (CTs) closer to CBDs are richer on average than those farther away as of 2000 across 16 US cities. This finding contradicts their theoretical results. However,

[1] In a first ring poor – second ring rich urban area a third ring of poor that use cars can emerge if the inner city does not have enough room for all the city's poor. In a rich first ring – poor second ring urban area a third ring of rich and a fourth ring of poor with car drivers could emerge.

this initial regression does not include CT proximity to rail transit. Once this dynamic is accounted for, being closer to the CBD is not as strongly correlated with income (see SI Fig. 1). The authors claim this is evidence of that PT infrastructure, generally concentrated in and near the CBD, helps explain pockets of poverty in the cores of American cities.

Second, using data from some of the US largest cities, they estimate the change in CT-level poverty rates in areas "treated" with new rail transit networks versus areas not treated with new rail transit networks. They find that treated CTs experienced greater increases in poverty rates after transit establishment than in CTs not impacted by new rail networks. Again, because these treatments tend to take place in denser urban areas near CBDs, the authors contend that these findings are consistent with modeled predictions of poorer HHs concentrating in an urban area's first ring in order to access their preferred commute technology.

Third, they show that in urban areas with little to no PT ($T_{Car} \approx T_{Rich} \approx T_{Poor}$) average HH income is higher near the CBD. Further, average HH income decreases with distance from the CBD. Conversely, in older cities with robust subway systems they find a U – shape relationship between HH income and distance from the CBD with the inflection point at 3 miles from the CBD. After this first ring of relatively rich HHs, poor HHs cluster to be near their preferred commute technology (approximately 3 miles from the CBD). In these cities, $T_{Car} \approx T_{PT}$ within a mile or two CBD. Both results are consistent with the basic monocentric city model with commute technology: when an urban area has only one commute technology or the urban area has two technologies but they are of equal speed (at least near the CBD) then a ring of rich HHs will emerge immediately around the CBD (inequality (1.4) does not hold).

1c. Shortcomings in Glaeser et al. (2008)'s analysis

While their analysis is impressive and thought-provoking, the Glaeser et al. (2008) analysis can be improved in several ways. I identify three major shortcomings. First, as mentioned above, Glaeser find values of T_{PT}, T_{Car}, and *F* that generate a modeled urban area consistent with observed HH income distribution pattern in the typical US city. However, they do not investigate whether *different types* of US urban areas have estimated values of T_{PT}, T_{Car}, and *F* that aligns modeled and observed patterns of HH income distribution (Fig. 3). For example, do cities like Boston and Chicago have T_{PT} and T_{Car} values that produce a rich inner ring in the monocentric city with commute technologies (MCCT) model? Further, do smaller urban areas have T_{PT} and T_{Car} values that align their observed income patterns with their modeled results?

Second, Glaeser et al. (2008) claim that "*[p]ublic transportation usage* appears to strongly predict poverty and to explain a substantial amount of the connection between proximity and poverty" (p. 15, emphasis mine). However, this claim is based on a regression that does not include *any* PT usage data. According to the authors, the tendency of the poor to co-locate with PT is enough to suggest that the poor rely more on PT than their rich counterparts. Later when they do conduct a regression analysis with PT usage data, correlations between CT-level income and CT-level usage of PT is enough for them to conclude that poor commuters are greater users of PT than rich commuters. Glaeser et al. (2018) never provide direct evidence of their claim that poor commuters use PT much more than the rich in most US cities or how commute mode use changes as commuters become poorer or richer. Nor do they attempt to clarify whether correlations between CT-level income and CT-level PT usage vary across different categories of US urban areas.

Third, Glaeser et al.'s event study of the impact of new rail networks on poverty in nearby CTs is flawed for several reasons. First, Glaeser et al. present no evidence that their event study was consistent with the common trend assumption needed for diff-in-diff identification. Second, their choice of control CTs – *all* CTs in select urban areas not treated with new rail networks – most likely does not allow for the causal inference they claim.

1.d. Re-evaluating the claims made in in Glaeser et al. (2008)

I re-evaluate the claims made in in Glaeser et al. (2008). I find the claim "*[p]ublic transportation usage* appears to strongly predict poverty and to explain a substantial amount of the connection between proximity and poverty" is not supported by data as of 2017. I argue that the word "strongly" in the above quote must be tempered to "somewhat" or a similar adjective.

My reasoning for a correction in tone is based on four analyses. First, I show that the MCCT model is not consistent with 2017 HH income spatial patterns in certain types of US urban areas when the model is parametrized with these areas' estimated 2017 T_{Car} and T_{PT}. In other words, something other than commute technology is affecting ring order in some US urban areas as of 2017.

Second, I show that Glaeser et al. (2008)'s claim $T_{Poor} \equiv T_{PT}$ and $T_{Rich} \equiv T_{Car}$ is not supported by 2017 data. In other words, Glaeser et al. (2008)'s back-of-the-envelope incentive compatibility constraint analysis that assigns poor commuters to PT use and rich commuters to car use is unreasonable in all American urban types. I show that 1) commuters from poor HHs overwhelming rely on cars for commuting across all US urban areas and 2) in the densest parts of the US's largest urban areas, rich commuters are just as reliant, if not more so, on PT than poor commuters.

Third, using a random utility model (RUM) of commute mode choice, I find that significant increases in a poor HH's income (approximately 50%) does not lead to a large increase in the probability that a poor commuter will switch from PT use to car use across all US urban types and rings within the urban types. This finding is inconsistent with Glaeser et al.'s claim that such income increase at a poor HH will lead to significant changes in commute mode choice (SI Text 2). Instead, I find that changes in distance to work and vehicle availability at the HH have a greater impact on probabilistic changes in mode choice than changes in HH income. In other words, almost all Americans, regardless of income, commute by car and changes in HH income, even significant ones, do little to change commute mode choice.

Finally, my event study analysis finding that new PT stations have little to no impact on neighboring HH incomes further corroborates my suggestion that PT location within an urban area has a more limited role on the spatial distribution of HH incomes than the one claimed by Glaeser et al. (2008). Using data from newly constructed light rail and express bus lines in Denver, Minneapolis / St. Paul, Los Angeles, and Phoenix I find very little evidence that areas immediately surrounding new stations experienced changes in income different than those of control group areas. I find my diff-in-diff inference more plausible than Glaeser et al (2008)'s event study analysis inference given my attention to the common trends assumption and deliberate construction of a control area set, two analytical steps that Glaeser et al (2008) do not take.

2. First analysis: Observed 2017 HH income spatial ordering in US urban areas generally cannot be explained by differences in commuting technology speeds and technology choices

Before I proceed to showing that estimated 2017 PT and car travel speeds are theoretically inconsistent with many observed spatial patterns of 2017 incomes within US urban areas, I first confirm that using the monocentric city model to analyze urban area structure as of 2017 is not grossly inappropriate. I do this by 1) showing the most urban area jobs are concentrated in or near the CBD and 2) that population density declines in distance from urban cores. Both spatial trends are assumed in the MCCT model.

In Fig. 4 I plot 2016 job density (at the zip code-level) against distance from the CBD across all metropolitan statistical areas (MSAs) with more than 1 million people and heavy rail PT (category 1) (Fig. 4A; see SI Table 1 for the MSAs that belong to category 1) and again across all MSAs with more than 1 million people and may or may not have light rail PT (category 2) (Fig. 4B; see SI Table 1 for the MSAs that belong to category 2).[2] According to the fitted splines, jobs in MSA categories 1 and 2 urban areas are most heavily concentrated in or near the CBD. I also confirm the MCCT model assumption of a monotonically decreasing population density gradient is evident in the MSA category 1 and 2 collections of urban areas as of 2017 (Fig. 5). I cannot create these plots for MSAs with less than 1 million people and no rail PT (category 3)

[2] Job density data for 2017 was not available at the time of manuscript preparation.

because the 2017 National Household Travel Survey (NHTS) (FHA 2017), the data source I use to categorize urban areas, does not identify which urban areas are in category 3. However, I have no reason to suspect that job and population density trends in category 3 urban areas are markedly different than those in categories 1 and 2. All in all, using the MCCT model to study the spatial pattern of HH incomes in America's urban areas does not appear to be grossly inappropriate.

To show estimated 2017 PT and car travel speeds are theoretically inconsistent with some observed spatial patterns of 2017 incomes within US urban areas, I first need to estimate inequality (1.4) parameters for the collection of urban areas in MSA categories 1, 2, and 3. Not only do I estimate T_{PT} and T_{Car} for each category of US urban area, unlike Glaeser et al. (2008), I also estimate a first and second-ring T_{PT} and T_{Car} for each urban area category. Fig. 5A suggests a first ring of 0 to 5 miles from the CBD and a second ring of 5 to 15 miles from the CBD in MSA category 1 urban areas. These choices are reasonable for two reasons. First, the average income gradient spline in MSA category 1 urban areas inflects around mile 5 (Fig. 5A). Second, the 2017 population density gradient spline intersects a density of 25,000 people per sq. mile at the 5-mile mark in MSA category 1 urban areas. Importantly, the 2017 NHTS indicates whether a respondent's home neighborhood has a population density of at least 25,000 per square mile or not. Therefore, it is not unreasonable to assign MSA category 1 survey respondents from the neighborhoods of 25,000 people per sq. mile or more to the inner ring of MSA category 1. I end the MSA category 1 second ring at 14 miles from the CBD because that is when the population density spline crosses the 10,000 people per square mile threshold, the next density threshold in the 2017 NHTS. Accordingly, I assign a MSA category 1 survey respondent from a home neighborhood with a population density in the 10,000 to 25,000 to MSA category 1's second ring. The appropriateness of this spatial splicing of MSA category 1 is corroborated by the finding that the average 2017 HH income among NHTS respondents assigned to ring 1 was \$102,441 (2017 USD) and income among NHTS respondents assigned to ring 2 was \$92,553 (2017 USD). In other words, the rings reflect a spatial hierarchy of average HH income, a feature of the MCCT model. In the case, MSA category 1 urban areas are consistent with a "rich" bid-rent curve that is steeper than "the less rich" bid-rent curve.

Further, Fig. 5B suggests a first ring of 0 to 10 miles from the CBD and a second ring of 10 to 23 miles from the CBD in MSA category 2 urban areas. Again, these choices are largely based on the data available from the 2017 NHTS. The MSA category 2 population density gradient spline never exceeds 10,000 people per square mile. After 10,000 people per square mile, the next density thresholds used in the 2017 NHTS are 4,000 and 2,000 people per square mile. Therefore, the NHTS respondent that lives in a MSA category 2 urban area CT with a population density of 4,000 or more is most likely to live in this urban type's first ring. Accordingly, I assign survey respondents from neighborhoods of 2000 to 4,000 people sq. mi.$^{-1}$ to a second ring of MSA category 2 urban areas. I assume similarly sized 1st and 2nd rings are found in MSA category 3 urban areas. In both of these cases, smaller average HH incomes in the first rings versus the second rings among NHTS respondents (\$79,943 versus \$91,448 in MSA category 2 urban areas and \$60,335 versus \$70,885 in MSA category 3 urban areas) is suggestive of a "less rich" bid-rent curve being steeper than "the richer" bid-rent curve.

I use 2017 NHTS survey data and the following model to estimate T_{PT} T_{Car}, and *F*,

$$Time_i = \alpha + \beta_1 \mathbf{1}[Car]_i + \beta_2 \mathbf{1}[Bus]_i + \beta_3 \mathbf{1}[Rail]_i + \beta_4 \mathbf{1}[Walk]_i + \beta_5 \mathbf{1}[Bike]_i$$

$$+\gamma_1 Dist_i + (\gamma_2 Dist_i \times \mathbf{1}[Car]_i) + (\gamma_3 Dist_i \times \mathbf{1}[Bus]_i)$$
$$+(\gamma_4 Dist_i \times \mathbf{1}[Rail]_i) + (\gamma_5 Dist_i \times \mathbf{1}[Walk]_i)$$
$$+(\gamma_6 Dist_i \times \mathbf{1}[Bike]_i) + \mathbf{FE}_i + \epsilon_i \quad i \in j \tag{2.1}$$

Unlike Glaeser et al. (2008), who estimated (2.1) across *all* US commuter survey data, I use data on commuters' neighborhood population density and urban area of residence to estimate (2.1) over nine sets of commuters (indexed by *j*): 1) commuters assigned to the first ring in MSA category 1 urban areas (MSA = 1, First) and in a similar fashion, commuters from 2) MSA = 1, Second; 3) MSA = 1, All respondents; 4) MSA = 2, First; 5) MSA = 2, Second; 6) MSA = 2, All respondents; 7) MSA = 3, First; 8) MSA = 3, Second; and 9) MSA = 3, All respondents.[3] Estimates of (2.1) over the commuter sets MSA = 1, All respondents; MSA = 2, All respondents; and MSA = 3, All respondents provide alternative estimates of T_{PT}, T_{Car}, and F for those readers who do not find my allocation of survey respondents into rings convincing.

In model (2.1), $Time_i$ indicates NHTS survey respondent *i*'s typical door-to-door commute time to work in minutes (i.e., the variable coded TIMETOWK in the survey), $\mathbf{1}[\quad]_i$ indicates *i*'s typical commute mode choice ('Other' is the omitted category), $Dist_i$ measures road network distance, in miles, between respondent *i*'s home and work (i.e., the variable coded DISTTOWK17 in the survey), and $\mathbf{FE}_i$ fixes respondent *i*'s urban location (e.g., Chicago CBSA versus LA CBSA).[4]

[3] Technically, Glaeser et al. estimate model (2.1) for each mode one at a time. Further, Glaeser et al. do not fix *i*'s urban location.
[4] When *i* is found in MSA category equals 1 or 2 $\mathbf{FE}_i$ indicates *i*'s home core-based statistical area (CBSA). When MSA category equals 3 $\mathbf{FE}_i$ fixes *i*'s home combination of US census division, MSA status, and presence of a subway system when population greater than 1 million (variable CDIVMSAR in the 2017 NHTS).

The sum of estimated $\hat{\gamma}_{1j} + \hat{\gamma}_{2j}$ is the estimate of $T_{Car,j}$ (minutes per mile commuting by car for group *j*) and the sum of estimated $(\hat{\gamma}_1 + \hat{\gamma}_3)BusW_j + (\hat{\gamma}_1 + \hat{\gamma}_4)RailW_j$ is the estimate of $T_{PT,j}$ (minutes per mile commuting by PT for group *j*). $BusW_j$ and $RailW_j$ are the person-weighted share of PT users that rely on bus and rail, respectively, for commuting in the j^{th} version of (2.1). Further, the estimated terms $\hat{\alpha}_{1j} + \hat{\beta}_{1j}$ and $(\hat{\alpha}_{1j} + \hat{\beta}_{2j})BusW_j + (\hat{\alpha}_{1j} + \hat{\beta}_{3j})RailW_j$ give the expected wait and egress time (in minutes) associated with car and PT use, respectively, for group *j*. For car users, wait and egress time includes the time needed to go from the parked car to work, and for PT users, it includes the time used to access the PT system and then access work once the PT system is left.[5] The wait and egress time for PT commuting relative to wait and egress time for car commuting for each group j – the parameter *F* in the inequalities (1.5)-(1.7) – is given by $(\hat{\alpha}_{1j} + \hat{\beta}_{2j})BusW_j + (\hat{\alpha}_{1j} + \hat{\beta}_{3j})RailW_j$ less $\hat{\alpha}_{1j} + \hat{\beta}_{1j}$.

As explained above, Glaeser et al. (2008)'s theoretical and empirical conclusions regarding the connections between the spatial allocation of PT and income within an urban area are also contingent on assigning $T_{PT,j}$ to poor commuters and $T_{Car,j}$ to rich commuters. They justify these assignments with the back of the envelope incentive compatibility analysis. However, because the 2017 NHTS includes each commuter's household income I can also directly estimate the differences in commute speeds and wait and egress time between poor and rich commuters. Therefore, unlike Glaeser et al. (2008), I also estimate the following model for each *j*,

[5] The wait and egress time for PT does NOT include the time spent transferring between PT vehicles. In the 2017 NHTS commuters who use PT are asked how much time they spend transferring among PT modes: "How many minutes each day do you usually spend transferring during your commute TO work (e.g. bus to bus, train to train, bus to train)?"

$$Time_i = \alpha + \beta_1 \mathbf{1}[Poor]_i + \beta_2 \mathbf{1}[Rich]_i + \gamma_1 Dist_i + (\gamma_2 Dist_i \times \mathbf{1}[Poor]_i)$$
$$+(\gamma_3 Dist_i \times \mathbf{1}[Rich]_i) + \mathbf{FE}_i + \epsilon_i \quad i \in j \tag{2.2}$$

where $\mathbf{1}[\quad]_i$ indicates *i*'s household income status. I assign a commuter from group *j* to the poor HH category if their household income is 138% or less of the 2017 federal poverty line (FPL) (I use this poverty threshold because members of households at 138% of the federal poverty line or lower are eligible for Medicaid). On the other hand, I assign a commuter from group *j* to the rich HH category if their household income is 400% or more of the 2017 FPL. Accordingly, the sum of $\hat{\gamma}_{1j} + \hat{\gamma}_{2j}$ is the estimate of $T_{Poor,j}$, the sum of $\hat{\gamma}_{1j} + \hat{\gamma}_{3j}$ is the estimate of $T_{Rich,j}$, and the estimate $\hat{\gamma}_{1j}$ is the estimate of $T_{Middle\ Class,j}$. Further, $\hat{\alpha}_j + \hat{\beta}_{1j}$, $\hat{\alpha}_j + \hat{\beta}_{2j}$, and $\hat{\alpha}_j$ are the estimated wait and egress of the typical poor, rich, and middle-class commuter from set *j*. If the poor of *j* overwhelmingly rely on PT for commuting and the rich of *j* overwhelmingly rely on car commuting then I should find that $T_{Poor,j} \approx T_{PT,j}$ and $T_{Rich,j} \approx T_{Car,j}$.

Finally, an empirical investigation of the MCCT model also requires estimates of Y_{Poor} and Y_{Rich} for each collection of urban areas. I set Y_{Poor} and Y_{Rich} for the collection of urban areas in MSA categories 1 and 2 equal to the lowest and highest points, respectively, on that category's median HH income spline (within the first two urban rings) (Fig. 6). I assume Y_{Poor} and Y_{Rich} numbers for MSA category 3 urban areas are equal to MSA category 2's numbers.

2a. Estimated commute technology speeds and commute technology choices generally do not generate monocentric city model bid-rent curves that are consistent with observed household income sorting across US urban areas

Estimates of model (2.1)'s parameters T_{Car}, T_{PT}, and *F* for the three US urban types and first and second rings in each urban area category are given in Table 1 (SI Text 3). Compared to the 3 minutes per mile national average Glaeser et al. (2008) found as of 2001, PT commute speeds in 2017 were faster on average in MSA categories 1 and 2 and slower on average in MSA category 3 urban areas. Further, compared to the 1.6 minutes per mile national average Glaeser et al. (2008) found as of 2001, car commuting speeds in 2017 were slower on average in MSA category 1 urban areas but much faster than average in the other two US urban area categories. Finally, compared to the 2001 national level estimate of 15 minutes, 2017 PT wait and egress time relative to car-based wait and egress were always lower no matter the urban type, ring category combination considered.

When I use MSA-wide estimates or ring-specific estimates of T_{Car} and T_{PT} and assume the typical poor commuter uses PT and the typical rich commuter uses a car (a la Glaeser et al. (2008)) the modeled MSA category 1 bid-rent curves are not consistent with the observation of a richer inner ring and a poorer second ring in these urban areas (Table 2). Only when poor second ring commuters are assumed to use *cars* and rich inner ring commuters are assumed to use *PT* do I find estimated MSA category 1 bid-rent curves consistent with the observed ring pattern. Conversely, estimated MSA category 2 and 3 urban area commute technology speeds and the assumption that the typical poor commuter uses PT and the typical rich commuter uses a car produces MCCT model bid-rent curves consistent with observed 2017 the ring patterns.

Notice my preliminary finding that estimated MCCT model bid-rent curves are consistent with the ring patterns observed across the various US urban area types as of 2017 relies on me assigning PT use to the rich and car use to the poor in MSA category 1 urban areas and vice-versa in MSA category 2 and 3 urban areas. However, based on two analyses, I find that such commute technology assignments are generally not supported by empirical data, with

MSA category 1 assignments being the exception. First, after I used model (2.2) to estimate $T_{Poor,j}$ and $T_{Rich,j}$ for each j (Table 3), I compared $\hat{T}_{PT,j}$ and $\hat{T}_{Car,j}$ to $\hat{T}_{Poor,j}$ and $\hat{T}_{Rich,j}$. In MSA category 2 and 3 urban areas I find that commute technology assignments that generate MCCT model bid-rent curves consistent with observed HH income patterns are not supported by comparisons of $\hat{T}_{PT,j}$ and $\hat{T}_{Car,j}$ to $\hat{T}_{Poor,j}$ and $\hat{T}_{Rich,j}$. To see this look at panels B and C of Fig. 7. As these panels make plain, there is little congruence between $\hat{T}_{PT,j}$ and $\hat{T}_{Poor,j}$ in MSA category 2 and 3 urban areas when using ring-specific or entire urban area estimates of T. However, $\hat{T}_{PT,j}$ and $\hat{T}_{Poor,j}$ need to be equivalent in these MSA categories if estimated MCCT model bid-rent curves are to be consistent with observed patterns. On the other hand, the congruence between $\hat{T}_{Car,j}$ and $\hat{T}_{Poor,j}$ and $\hat{T}_{Car,j}$ and $\hat{T}_{Rich,j}$ in MSA category 2 and 3 urban areas is striking. In other words, poor and rich commuters both overwhelmingly commute by car in these US urban areas.

In contrast, the estimates of $\hat{T}_{PT,j}$ and $\hat{T}_{Car,j}$ and $\hat{T}_{Poor,j}$ and $\hat{T}_{Rich,j}$ in the first two rings of MSA category 1 urban areas support the commute technology assignments that generate MCCT model bid-rent curve consistent with observed HH income patterns (Panel A of Fig. 7): $\hat{T}_{Car,j}$ falls within $\hat{T}_{Poor,j}$ and $\hat{T}_{PT,j}$ falls within $\hat{T}_{Rich,j}$ (only the urban-wide estimates of $\hat{T}_{Rich}$ and $\hat{T}_{PT}$ do not line up). In other words, only in the first two rings of MSA category 1 urban areas do I find all of the empirical conditions consistent with MCCT model equilibrium, albeit in a manner inconsistent with Glaeser et al. (2008)'s overall thesis where the poor rely on PT and the rich on cars.

The irrelevance of PT to almost all commuters, poor and rich alike, in MSA 2 and 3 urban areas is corroborated by the observed distribution of commute mode choices in the 2017 NHTS

(Table 4). Overall, approximately 10 percent of commuters from poor HHs in these urban areas use PT to commute to work. Even inner ring poor commuters in these urban areas, the group in these areas most likely to use PT according to Glaeser et al. (2008)'s thesis, only use PT at a rate slightly greater than 10%. Commuters from middle-income and rich HHs are even less likely to use PT to commute to work in these urban areas.

Conversely, in MSA category 1 urban areas a rich commuter from ring 1 is more likely to use PT than a poor commuter from ring 2 (Table 4). Further, in the first ring of MSA category 1 urban areas PT share is greater than 50% for both poor and rich commuters. Further, consistent with estimated model results that support a MCCT model equilibrium in MSA category 1 urban areas, most poor commuters from the second ring in MSA category 1 urban areas use a car rather than PT (interestingly, the ring 2 poor use PT less than the ring 2 rich). Therefore, the 2017 NHTS data suggests that PT is a highly relevant commute mode in America's largest urban areas with rail and that HH location decisions in these urban areas are likely affected by PT access or at least that commute mode choice is a function of residential location. However, there is little evidence to suggest that poor HHs are overwhelmingly attracted to and rich HHs are largely repelled by PT-dense areas. If anything, it appears that rich commuters are just as attracted to PT, if not more, than poor commuters in the US' largest urban communities serviced by rail (see SI Text 4 and SI Table 2 for an analysis of 2017 commuting patterns using American Community Survey (ACS) data).

2b. Summary of the first analysis

To summarize my first re-analysis of Glaeser et al.'s (2008) work, I find that estimated commutes technology speeds in the US' largest cities with heavy rail PT (category 1) as of 2017 produce theoretical poor and rich HH bid-rent curves that are generally consistent with the observed income patterns and commute mode choices in these cities as of 2017. However, for this result to hold richer commuters that live near the CBD must be strong users of PT and poorer commuters that live further from the CBD need to be willing and able to use cars for commuting. My analysis of NHTS data suggests these commute choice patterns existed in the US' largest cities with rail heavy PT as of 2017. However, the empirical evidence that poor commuters in the first rings of 1) large US cities without heavy rail (category 2) and 2) smaller US cities (category 3) overwhelmingly chose car commute technology means that the "typical" US city cannot be explained by a MCCT model equilibrium where the poor commuters near the CBD take the slower PT and richer commuters further out rely on cars.

All in all, while PT networks may explain the pattern of HH income rings in some of the US' most iconic cities (New York, Chicago, etc.), I believe it is a stretch to suggest that income patterns within America's more typical cities can be explained by PT networks. Empirical evidence suggests that the use of PT commuting technology in these cities is too insignificant to make much of a difference in the broader income distribution patterns.

3. Second analysis: A commute mode choice model does not suggest that poorer commuters significantly increase their car commuting rates after a significant increase in HH income, at least in the short run

Glaeser et al. (2008)'s thesis is that mode choice is a predominately a function of income. In most US cities, they argue, poor commuters have financial incentive to choose PT and that once a poor HH becomes significantly wealthier it will find it rational to switch to car commuting.[6] I assess this assertion by parametrizing a random utility model (RUM) of modal choice with the 2017 NHTS respondents. (RUMs have often been used to estimate commute mode choice e.g., McFadden (1977), Bhatta and Larsen (2011); Cartenì et al. (2016).) I find that poor commuters do not flock to car commuting given a significant increase in income. Instead, other changes in HH characteristics do a better job of explaining mode choice changes among poorer commuters.

I assume that a commuter chooses transportation mode *l* over all other modes, indexed by *j* = 1,…,*J*, if that choice maximizes their utility from commuting,

$$V_l + \varepsilon_l > V_j + \varepsilon_j \; \forall j \neq l \tag{3.1}$$

where $U_l = V_l + \varepsilon_l$ is the commuter's utility from choosing *l*, V_l is a function of observable covariates, and ε_l is a function of unobservable covariates, only known to the commuter. Therefore, the probability that the commuter will choose mode *l* over all other modes is,

$$\Pr\left(\varepsilon_1 < V_l - V_1 + \varepsilon_l\,, \varepsilon_2 < V_l - V_2 + \varepsilon_l, \dots, \varepsilon_J < V_l - V_J + \varepsilon_l\right) \tag{3.2}$$

The known part of (3.1) for commuter *i* is given by,

$$V_{ij} = \alpha_j + \beta \mathbf{x}_{ij} + \delta_j \boldsymbol{\omega}_{ij} + \gamma \mathbf{z}_i + \theta \mathbf{t}_j \tag{3.3}$$

The first term of (3.3) is the model intercept for mode choice *j*. The vectors $\mathbf{x}_{ij}$ and $\boldsymbol{\omega}_{ij}$ both contain mode-commuter specific covariates such as commuter *i*'s time to work using mode *j* (T_{ij}) and commuter *i*'s cost of using *j* (C_{ij}). If a mode-commuter variable's effect on *V* differs across *j* then it is part of vector $\boldsymbol{\omega}_{ij}$; otherwise it is part of vector $\mathbf{x}_{ij}$. For example, if a minute spent commuting on a bus effects *i*'s utility differently than a minute spent commuting in a

[6] "These [estimates for *F*, *C*, W_{Rich}, and W_{Poor}] also suggest that a[n]…increase in income from \$10 to \$20 per hour should be associated with a massive shift from public transportation to driving." (p. 13).

private car then T_{ij} is part of $\boldsymbol{\omega}_{ij}$. The vector $\mathbf{z}_i$ contains commuter-level variables such as income, gender, age, etc. Finally, the vector $\mathbf{t}_j$ contains mode-level variables.

The only mode-commuter covariate in the 2017 NHTS is T_{ij}. However, T_{ij} is only observed for the mode that commuter *i* chooses. Further, there are no mode-level variables in the 2017 NHTS. Therefore, my default specification of equation (4.3) only uses commuter-level variables.

$$V_{ij} = \alpha_j + \gamma_1 I_i + \gamma_2 D_i + \gamma_3 A_i + \gamma_4 M_i + \gamma_5 W_i + \gamma_6 V_i \quad (3.4)$$

The variable I am most interested in, commuter *i*'s annual HH income, is given by I_i. The impact that HH income has on mode choice is not consistent across the mode choice literature. For example, Shen et al. (2016) found that income was positively associated with commuting by car in four Shanghai suburban neighborhoods (also see Bhat and Sardesai 2006). On other hand, a 1998 travel survey in the Netherlands found that household income had little effect on commute mode choice (Limtanakool et al. 2006).

The variable D_i measures *i*'s distance to their place of work. D_i is positively correlated with T_{ij} (see Table 6), one of the mode-commuter variables I would prefer to use in lieu of D_i in model (4.4). Presumably D_i is also positively correlated with C_{ij}, another mode-commuter variable I would prefer to use in model (4.4). Given that past research has found that a small increase in PT's *C* and *T* decreases a commuter's utility more than a small increase in car's *C* and *T* (Frank et al. 2008, Limtanakool et al. 2006), I expect increases in D_i will reduce and increase the probability of PT use and car use, respectively, all else equal.

I include the gender of the commuter (M_i = 1 if *i* is a man and equals 0 otherwise) in model (3.4) because past researchers have found men are more likely to use a car for commuting than women, all else equal (Limtanakool et al. 2006). Speculated reasons for this divergence include inequality in monetary rewards from working, the spatial distribution of jobs and household task allocation, and women's weaker bargaining power in household scarce allocation decisions (Limtanakool et al. 2006, Scheiner and Holz-Rau 2012).

The number of vehicles per driver in the household (V_i) controls for i's ability to access a car for commuting. Of course, car commuting is impossible if the household does not own one. Further, Bhat and Sardesai (2006) and Limtanakool et al. (2006) found individuals in households with a high number of vehicles per licensed adult are more likely to choose car commuting given there is less competition for cars. Other variables in model (3.4) include i's age (A_i) and race (W_i = 1 if i is a white person and equals 0 otherwise). I include age because I assume a commuter's age may affect their willingness or ability to choose each mode j (Eluru et al. 2012).

Alternatively, I can use multiple imputation (MI) methods to impute the missing T_{ij} values when $j \neq l$. In the modified RUM that used imputed T_{ij} I would drop D_i given the high level of correlation between the two variables.

$$V_{ij} = \alpha_j + \delta_j T_{ij} + \gamma_2 I_i + \gamma_3 A_i + \gamma_4 M_i + \gamma_5 W_i + \gamma_6 V_i \tag{3.5}$$

where T_{ij} is mode-commuter variable that effects V differently depending on j (i.e., the variable is part of vector $\boldsymbol{\omega}_{ij}$ not $\mathbf{x}_{ij}$).

3a. Imputing commute time

Multiple imputation (MI) can generate estimates of missing travel times that are valid for population-level inference if travel times are missing at random (MAR). These data are MAR if the probability of their "missing"-ness does not depend on unobserved data but rather can be explained by data recorded in the 2017 NHTS. Note that the pattern of missing times is completely explained by commuters' mode choices. Therefore, given my RUM analysis assumes that 2017 NHTS covariates can largely explain mode choice, I am assuming, by definition, that the pattern of missing commutes times can also be largely explained by 2017 NHTS covariates. Given there is no test for MAR, an assumption of MAR will have to do.

For a continuous variable with a restricted range, such as travel time, nearest matching (Raghunathan et al. 2001) or 'predictive mean matching' (PMM) is a recommended MI method. To implement PMM I first divided the 2017 NHTS into separate MSA category – mode choice datasets. For example, the "MSA category 1 – car" dataset includes every MSA category 1 commuter, their responses to survey questions, and if they commuted by car to work in 2017, their travel time. For those that did not commute by car their travel time in the "MSA category 1 – car" dataset is unobserved. Second, I than ran the PMM method over each dataset ten times to obtain ten complete sets of T for each i where T_{il} is always equal to i's observed travel time and T_{ij} for all $j \neq l$ is imputed in each set. PMM imputes T_{ij} when it is unobserved ($j \neq l$ for i) by setting it equal to the travel time for one of its nearest 5 commuter "neighbors" that has an observed travel time for mode j (SI Text 5). In each MI iteration the neighbor that 'donates' their T_{ij} is randomly chosen (SI Text 3).

3b. Estimating the RUM models

I assume the multinomial logit functional form where the dependent variable is equal to 1 if i typically used mode $j = l$ and equal to 0 for all other j when estimating (3.4) and (3.5).[7] I

[7] Alternatively, a nested logit can be used to estimate these models when the assumption of the independence of irrelevant alternatives (IIA) is presumed not to hold between all mode choices. The IIA property may not hold in

generate estimates of (3.4) and (3.5) for each income class (poor or rich) – ring (1, 2, or entire urban area) – MSA category (1, 2, or 3) combination (see SI Text 3 for instructions on using R to replicate all multinomial logit estimates of models (3.4) and (3.5)). Further, I use the person weights assigned to each commuter when estimating (3.4) and (3.5). Because each commuter has 10 sets of T_{ij} there are 10 estimates of (3.5) for each income class – ring – MSA category.

To make interpretation of model estimates easier to understand I use standardized continuous variables when estimating (3.4) and (3.5). The variables are standardized according to the range of observations in each geographic subset of the data. For example, data in the poor commuter – ring 1 – MSA category 1 estimates of (3.4) and (3.5) are standardized according to the means and standard deviations in the poor commuter – ring 1 – MSA category 1 dataset (see Table 5 for the mean and standard deviation of each independent variable in models (3.4) and (3.5) for poor and rich commuters across each geographic subset).

An estimated function (3.4) can be used to calculate the probability that the mean poor or rich commuter from a certain type of urban area, living in its first or second ring or the area generally, chose mode j in 2017,

$$\hat{P}_{j,Inc,Geog,MSA} = \frac{e^{\alpha_j+\hat{\gamma}_1\bar{D}+\hat{\gamma}_2\bar{I}+\hat{\gamma}_3\bar{A}+\hat{\gamma}_4\bar{M}+\hat{\gamma}_5\bar{W}+\hat{\gamma}_6\bar{V}+\hat{\gamma}_7\overline{ST}+\hat{\gamma}_8\overline{Manu}+\gamma_9\overline{FT}}}{\sum_{j=1}^{J} e^{\alpha_j+\hat{\gamma}_1\bar{D}+\hat{\gamma}_2\bar{I}+\hat{\gamma}_3\bar{A}+\hat{\gamma}_4\bar{M}+\hat{\gamma}_5\bar{W}+\hat{\gamma}_6\bar{V}+\hat{\gamma}_7\overline{ST}+\hat{\gamma}_8\overline{Manu}+\gamma_9\overline{FT}}} \tag{3.6}$$

where the bar above each variable indicates the mean of the standardized variable and the subscript j,*Inc*,*Geog*,*MSA* indicates the mode choice j, income category, geography (a ring or the

commute mode choice models because several modes, including the bus and rail choices and the walk and bike choices, may be treated as substitutes by many commuters (Shen et al. 2016). In a nested logit estimate of (3.4) and (3.5) I would use "nests" of mode choices such that IIA holds within each nest of alternatives but not across nests (Train 2003, Shen et al. 2016). For example, I could treat ''bus and rail transit" and ''bike and walk" as two nests of alternatives.

urban area in general), and urban area type combination. Further, the estimated function (3.5) can be used to calculate the probability that the mean poor or rich commuter from a certain type of urban area, living in its first or second ring or the area generally, chose mode *j* in 2017 when the independent variable T_{ij} is used in lieu of D_i,

$$\hat{p}_{j,Inc,Geog,MSA} = \frac{e^{\alpha_j+\hat{\delta}_j\bar{T}_j+\hat{\gamma}_2\bar{I}+\hat{\gamma}_3\bar{A}+\hat{\gamma}_4\bar{M}+\hat{\gamma}_5\bar{W}+\hat{\gamma}_6\bar{V}+\hat{\gamma}_7\overline{ST}+\hat{\gamma}_8\overline{Manu}+\gamma_9\overline{FT}}}{\sum_{j=1}^{J} e^{\alpha_j+\hat{\delta}_j\bar{T}_j+\hat{\gamma}_2\bar{I}+\hat{\gamma}_3\bar{A}+\hat{\gamma}_4\bar{M}+\hat{\gamma}_5\bar{W}+\hat{\gamma}_6\bar{V}+\hat{\gamma}_7\overline{ST}+\hat{\gamma}_8\overline{Manu}+\gamma_9\overline{FT}}} \tag{3.7}$$

Given there are 10 estimates of model (3.5), due to 10 unique set of T_{ij} across all *i*,*j*, $\hat{p}_{j,Inc,Geog,MSA}$ is calculated 10 times for each instance of *j*,*Inc*,*Geog*,*MSA*.

3c. Simulating the impact of a small change in each covariate on mode choice probability

Consistent with overall patterns I found in section 1 of this paper, the first rings of MSA category 1 urban areas were the only places in urban America where the car was not the dominate commute mode in 2017 (see Table 6 for $\hat{P}_{j,Inc,Geog,MSA}$ across all *j*,*Inc*,*Geog*,*MSA* combinations except *j* = other). Further, within these densest US urban areas the mean rich commuter was more likely to use rail PT than the average poor commuter, while the average poor commuter was more likely to use bus PT than the average rich commuter. Otherwise, throughout the rest of urban America, the mean rich commuter is 8 to 16 probability points more likely to use a car than the average poor commuter (see SI Text 3 for instructions on using R to find $\hat{P}_{j,Inc,Geog,MSA}$)

Next, I simulated the *ceteris paribus* effects of 1 standard deviation (SD) changes in a mean commuter's explanatory variables on their mode choice probabilities (or in the case of a dummy variable, a change in the variable's binary status). Across all commuter types, an increase in distance to work (*D*) and vehicle availability (*V*) is most likely to lead to a change in commute mode technology (Fig. 8 and SI Table 3). In general, representative commuters that suddenly had further to travel or gained greater access to a car and had bused, trained, walked, or biked to work became more inclined to use a car. The one exception to this rule could be found in MSA category 1 urban areas where an increase in distance to work encouraged greater rail PT use as well (see SI Text 3 for instructions for running the R script that generates simulated numbers).

Another striking result from the simulated changes in $\hat{P}_{j,Inc,Geog,MSA}$ are the differences in mode-switching elasticities across income classes and urban classification categories. First, the mean poor commuter's mode choice probabilities, rather than the mean rich commuter's probabilities, changed the most in response to exogenous changes in independent variables. Second, the mean first ring commuter's mode choice probabilities, rather than the mean second ring commuter's probabilities, changed the most in response to exogenous changes in independent variables. Third, the mean MSA category 1 commuter's mode choice probabilities, rather than the mean MSA category 2 or 3 commuter's probabilities, changed the most in response to exogenous changes in independent variables. I suspect the latter two patterns can be explained by 1) the greater density of PT, sidewalks, and bike lanes in the US's most densely populated areas (Newman and Kenworthy 1989 and Kenworthy and Laube 1999, Frank and Pivo 1994, Glazier et al. 2014, Whalen et al. 2013, El-Assi et al. 2017, Schoner and Levinson 2014) and 2) the greater costs of owning a car in the inner city relative to the suburbs (parking may be difficult or, if available, expensive and auto theft and vandalism are much more prevalent in the inner-city). Both of these dynamics make it easier for inner-city commuters to avoid or abandon car commuting if circumstances change relative to commuters that live in less dense areas with lower levels of PT and infrastructure that encourages biking and walking.

Finally, contradicting Glaeser et al. (2008)'s claim that a substantial increase in a poor commuter's income would lead to a sharp intake of car commuting, I find that a 1 SD change in the mean poor commuter's HH income (approximately a 55% increase in annual HH income) has little effect on their expected mode choice probabilities. The greatest change in the probability of using a car after a 1 SD increase in HH income was 5 percentage points across all the various poor commuter – urban geography – urban type combinations. Of course, the relatively weak relationships between increases in HH income and switching to car use among poor commuters could be explained by the vehicle per driver variable (*V*) in (2.4). If relatively car-scarce poor HHs tend to use significant increases in income to buy cars then increases in income at a poor HH *would* make car commuting by the HH's workers substantially more likely (recall that a 1 SD increase in *V* is associated with a 10 to 20 percent probability point increase in the mean poor commuter using a car to commute to work).

To empirically test if a 1 SD in *V* is masking the effect of an increase in HH income on commute mode choice among the poor I re-estimate (2.4) and then re-simulate the impact of a 1 SD change in each of the average commuter's independent variables on their mode choice probabilities *after* dropping *V* from (2.4). Generally, the probability of a mean poor commuter switching to car commuting increases by an additional 5 percentage points after a 1 SD increase in their HH's income compared to simulations when *V* is included (SI Text 6). In other words, even when a ~50% increase in HH income is the only channel to increasing car availability at a poor HH, the mean poor commuter's shift car use only increases approximately 10 percentage points; nowhere near as dramatic an increase as predicted by Glaeser et al. (2008). (Further, see SI Text 6 and SI Table 5 for a discussion on why the additional 5 percentage point increase in car use among poor commuters per 1 SD increase in HH income after *V* is dropped from model (2.4) is likely an overestimate due to omitted variable bias.)

Replacing less informative distance to work with the more information-rich time to work by mode in the mode choice model made little difference in modeling results (Table 7, Fig. 9, SI Text 3, and SI Table 4). Commuters in the largest urban areas, commuters in the first ring of a city, and poorer commuters much more readily changed modes given an exogenous change in their conditions than commuters in smaller cities, commuters in the second ring of a city, and richer commuters. Further, I still find that a 1 SD increase in vehicle ownership at a poor HH is much consequential than a 1 SD increase in income at a poor HH when it comes to increasing car commuting among these HHs. Specifically, the mode choice model that includes travel time to work in lieu of distance to work indicates that a 1 SD increase in *V* has at least 3-times the effect on car choice probabilities than a 1 SD increase in HH income across all HH income – location – urban type combinations.

Further, when I dropped *V* from (2.5) and then re-simulated changes in $\hat{p}_{Car,Poor,Ring,MSA}$, I once again find the probability of a mean poor commuter using a car for commuting increases by an additional 5 percentage points after a 1 SD increase in their HH's income compared to model estimates and simulations where *V* is included in (2.5) (SI Table 4). Despite this boost, the $\Delta\hat{p}_{Car,Poor,Ring,MSA}$ associated with a 1 SD increase in *V* in the version of (2.5) that includes HH income is still larger than the $\Delta\hat{p}_{Car,Poor,Ring,MSA}$ associated with a 1 SD increase in HH income when an increase in income is the only channel to increasing car availability at a HH. (And again, due to omitted variable bias in (2.5), the estimated $\Delta\hat{p}_{Car,Poor,Geog,MSA}$ per 1 SD increase in HH income among poor commuters when *V* is dropped from (2.5) is likely an overestimate. See SI Text 6 and SI Table 5.)

3c. Summary of the second analysis

While poor commuters do increase their rate of car commuting as they get significantly richer (I simulate a 50% or so increase in the mean poor HH's income). The change in probability is nowhere near as dramatic as suggested by Glaeser et al. (2008)'s incentive compatibility analysis. Other commuting-related variables, including changes in the availability of vehicles at the HH, changes in distance or time to work, and changes in age or gender are just as likely or even more likely to explain changes in poor commuter mode behavior. Overall, in the short run at least, commuter mode choice reactions to exogenous shocks appear to be small no matter their HH income. Given the evidence in section 2 that most commuters already rely on cars for commuting no matter their income, finding relatively small changes in commute mode choice with respect to any exogenous shock across the HH income spectrum is not surprising: there are relatively few American commuters that have not already switched to car use.

The densest neighborhoods of the largest cities in the US (MSA category 1) are the only places in the US where my claim of little mode switching due to exogenous shock can be challenged. Further, poor commuters in these areas more readily switch commute modes in response to exogenous shocks than their richer neighbors, at least in the short run. However, poor commuters' willingness to switch to PT use in these areas appears to largely be a function of incidental access to PT, not a strong preference for PT or for locating to areas with dense PT networks. I base this last conclusion of two pieces of evidence. First, in the parts of urban America *not* located in the first rings of MSA category 1 urban areas, poor commuters rarely switch to PT use no matter the exogenous shock. Second, as I will show in the next section, I find little evidence to support the claim that poor HHs tend to cluster in neighborhoods with growing PT access.

One note of caution about the mode choice elasticities I have estimated here. These are short-run responses by commuters to income, distance to work, time to work, and other shocks. In other words, the modeled and simulation results assume that HHs remain in their geographic area after experiencing a commute-affecting exogenous shock. In the long-run, commuters that experience permanent changes in income, distance to work, or other commuting related factors may relocate to maximize their HH's utility. For example, after a job change causes distance to work to increase, a poor commuter may eventually move their HH closer to their job. Or, in the long-run, poorer HHs that become richer and lived in the first ring of their city may move to the suburbs, thereby becoming even more reliant on cars for commuting than indicated by the model results discussed above. Conversely, HHs that lost income could eventually move within a city to improve their access to bus routes. Relocation by HHs because of exogenous shock leads to commuting decisions beyond the scope of my analysis.

4. Third analysis: There is little evidence that urban neighborhoods treated with new PT infrastructure in the mid-2000s disproportionally attracted poor households

In my simulations of mode choice responses to exogenous shocks I also found that poor workers in MSA category 2 and 3 urban areas were slightly more likely to use a bus rather than a car after their HH's incomes fell, all else equal (see Figs. 8 and 9). Based on these results one could surmise that some poor commuters who have experienced a drop in their HH incomes drop would eventually move closer to PT networks for better access to a commute mode they increasingly find more appealing.

A trend of poorer HHs concentrating closer to PT networks over time is important to Glaeser et al (2008).'s overall hypothesis. And, in fact, Glaeser et al. (2008) claim to have found evidence of this dynamic. Using data from 16 US urban areas, Glaeser et al. (2008) argue their finding that CTs "treated" with new transit lines between 1980 and 2000 subsequently became more impoverished (the poverty rate increased) relative to CTs that did not gain access to new transit lines during this same time period (see section 5 of Glaeser et al. (2008)) is evidence that PT infrastructure *caused* nearby areas to become poorer.

However, there are reasons to doubt their causal claim. Their claim that new transit *caused* surrounding neighborhoods to become poorer would be defensible if they showed that 1) their choice of control areas allowed for causal identification; 2) the average difference in HH incomes between CTs 'treated' with new PT infrastructure and 'control' CTs was constant over time prior to treatment (i.e., the common trend assumption); and 3) the average difference in HH incomes between the 'treated' and 'control' areas were different post-treatment. While they find the post-treatment effect (albeit, of very small magnitude), Glaeser et al. (2008) present no evidence that their event-study used a control set consistent with causal identification and that it satisfied the common trend assumption.

While their event-study results may have incidentally satisfied the common trends assumption (they do not provide the data and code resources to verify this), their control set choices make claims of causal identification unlikely. Their control set includes *all* CTs within their 16 study urban areas not within one mile of rail transit established between 1980 and 2000. Therefore, their control set includes areas of the cityscape that could have plausibly experienced PT expansion between 1980 and 2000 *and* areas that could not have plausibly experienced PT expansion in the event-study time window (e.g., geological features or zoning restrictions made such expansion impossible). However, the use of a control group that includes both types of areas violates the basis of identification in event-study designs. An event-study's causal identification relies on important unobserved variables – omitted factors that can help explain the expansion of PT in my case – as either being time-invariant group attributes (e.g., *both* treated and control areas are in major transportation corridors and could have plausibly been the site of expanded PT) or time-varying factors that are group invariant (e.g., *both* treated and control areas experience zoning changes over time that made PT expansion in both sets of areas increasingly more likely) (Wing et al. 2018). Indiscriminately selecting *all* untreated CTs for the control set surely violates both necessary conditions for claiming causality in an event-study.

Beyond using a control set that makes claims of causality unsupportable, a second reason to question Glaeser et al. (2008)'s event-study conclusion is illuminated by recent work showing event-study results are likely biased if the analysis includes treated observations that subsequently become controls (e.g., Abraham and Sun 2018, Goodman-Bacon 2019). For example, suppose 10 areas in an urban area were treated with PT in 2000, 20 in 2005, and 10 in

2010. Further, suppose I ran an event-study where each of these areas is assigned the treatment indicator variable at the time of treatment and beyond. Under this framework, the set of control observations changes over time (e.g., the areas treated in 2005 are compared to the union of the original control set and the areas treated in 2000 while the areas treated in 2000 are only compared to the original control set). According to Abraham and Sun (2018) and Goodman-Bacon (2019) this event-study's results would likely be biased. In the Glaeser et al. (2008) event-study (it appears) the roster of the control set changes over time and yet they do not attempt to correct for the potentially bias results this approach can create (I say appears because they do not provide the data and code resources to verify that the control set changes over time in their event-study).

4a. My strategy for determining if new PT stations make the areas near to the station poorer or richer relative to other neighborhoods in the same city

In the event-studies I use to measure the impact of new PT infrastructure on average incomes in nearby areas I 1) meet the necessary conditions for claiming causality by either using a set of control areas entirely comprised of plausible candidates for transit expansion or a set of control areas entirely comprised of non-plausible candidates for transit expansion during the event study window; 2) consider and evaluate pre-trends; and 3) use a roster of control areas that does not change over time. In addition, I use income data defined at a finer spatial grain than that used in the Glaeser et al. (2008) study. Finally, I verify that my findings are robust to several measures of neighborhood income and not just a pattern found in one measure of income a la Glaeser et al. (2008).

In my event-studies "treated" areas are the one-half mile radii areas (0.79 square miles) encircling light-rail or rapid-bus stations that opened in year $\underline{t} < t < \bar{t}$ in urban area *u* be where $\underline{t}$ and $\bar{t}$ indicate the beginning and end, respectively, of the event-study window. I use a half-mile radius treatment area because US transit agencies generally assume PT station influence extends in all directions for a half-mile.[8] When necessary, I truncated treated areas to avoid overlapping.[9]

The comparative control areas in each event-study analysis I run are either comprised of 1) nonoverlapping half-mile radius circles centered on randomly chosen points along street car or trolley routes that existed in *u* prior to 1960 but were not treated by light-rail or rapid-bus stations between $\underline{t}$ and $\bar{t}$, 2) the 0.5 mile-width buffer immediately surrounding treated areas[10], or 3) nonoverlapping half-mile radius circles centered on light-rail or rapid-bus stations that are slated to open in *u* after $\bar{t}$. I ensure that treated areas and control areas never overlap.

The first set of controls – areas on a city's historical streetcar or trolley network that subsequently did not host a light-rail or rapid-bus station between $\underline{t}$ and $\bar{t}$ – are causal identification-appropriate treated area matches for two reasons. First, treated areas in my event-study cities of LA, Denver, and Minneapolis align with these cities' historical networks as well, suggesting that the control areas along the historic network could have been viable

[8] See the Federal Transportation Administration's Reporting Instructions for the Section 5309 Capital Investment Grant Program (https://gisdata.mn.gov/dataset/us-mn-state-metc-trans-station-areas-half-mile). Hurst and West (2014) use a half-mile radius impact zone in their analysis of new light rail stations' effect on land use change in Minneapolis.

[9] Treated areas are found by 1) creating a Thiessen polygon map where polygons are centered on the stations in the focal PT network, 2) generating half-mile radius circles around each station in the focal network, and 3) clipping the Thiessen polygons with the circle polygons.

[10] These buffer areas are clipped such that it includes no areas "treated" by existing or future light-rail or rapid-bus transit stations. Unlike the nonoverlapping half-mile radius circles centered on new PT stations, PT stations that will open after $\bar{t}$, and points along historical street car or trolley routes, these control buffer areas can overlap, however.

candidates for stations established between $\underline{t}$ and $\bar{t}$ (common time-invariant group attributes). Second, Brooks and Lutz (2019) have shown that areas around early 20th-centruy LA streetcar stations have experienced similar population and building density and zoning trends over the last 100 years despite the heterogeneity in exogenous shocks these areas have experienced since then. Assuming these trends hold in other cities, we can surmise that unmeasured variable trends in *u*'s treated areas are very similar to those in control areas along the historical network up to treatment year $\underline{t} < \tau < \bar{t}$. Therefore, an assumption of common socioeconomic trends between treated and historical-transit control areas prior to treatment is not unreasonable.

The second set of controls I use – the 0.5 mile-width buffer immediately surrounding treated areas – are areas that could have plausibly been treated between $[\underline{t}, \bar{t}]$ as well but were not and likely experienced the same general density, zoning development, and socioeconomic trends as the treated areas prior to treatment year τ. For example, distance to urban amenities and jobs will be similar across a treated area and its control. Further, socioeconomic and economic attributes that tend to spatially cluster in cities – such as race, income, and housing stock age – will be very similar across both a treated area and its control buffer.

Common across the treated and control sets in the final set of controls I use – one-half mile radii circles around stations slated to open after $\bar{t}$ – is the regional planner's judgment that both sets of areas are suitable for PT network expansion. Therefore, a common trend assumption in the unobserved transit location choice-related variables between treated areas and these control areas is reasonable. However, because treatment in control areas before $\bar{t}$ was not plausible given a post-$\bar{t}$ construction date, I should be able to identify differences in income patterns across the two sets of areas due to treatment. Crucially, I assume that people and businesses have not yet incorporated the opening of future stations into their decision making as of time $\bar{t}$. This is essentially the assumption of a "no anticipation condition" that I discuss below.

The model I used to identify the average effect of a PT station's establishment on its surrounding area income trends relative to average income trends in control areas is given by,

$$M_{uct} = \alpha + \theta_{uc}\mathbf{1}[Treat]_{uc} + \sum_{t=\underline{t}}^{\bar{t}} \beta_{ut} D_{uct}^{\tau}$$

$$+ \sum_{t=\underline{t}}^{\bar{t}} \delta_{tu} D_{uct}^{\tau}\, \mathbf{1}[Treat]_{c} + \mu_{uc} + \epsilon_{uct} \qquad (4.1)$$

where M_{uct} is the average median household income, per capita income, or poverty rate in area *c* in urban area *u* in year *t* (2017 USD), D_{uct}^{τ} equals 1 if area *c* in *u* is $t - \tau$ years before (if $t - \tau$ is negative) or years after (if $t - \tau$ is positive) treatment year τ and equals 0 otherwise, $\mathbf{1}[Treat]_{uc}$ is equal to 1 if *c* is a treated area and equal to 0 if not, and μ_{uc} is area *c*'s fixed effect. In each estimate of (4.1) I only include areas that were treated in the same year (given by τ), thereby avoiding the bias that can be created when treated areas eventually become control areas in an event-study. Therefore, D_{uct}^{τ} is the same for all *c* in *u* across all *t*.

I estimated each unique event-study analysis twice, once using M_{uct} created with US Census' block-group (BG) data and again with CT-level income data. To calculate a BG and CT-level M_{uct} for each unique {*u*,*c*,*t*}-tuple I first converted *u*'s BG or CT-level median HH income, per capita income, or poverty rate polygon map from year *t* (Manson et al. 2019) into a 10 x 10-meter grid cell map (each cell has an income level equal to its parent polygon). I then found the average median HH income, per capita income, or individual poverty rate across the grid cells contained in area *c*'s polygon.

BGs are nested within CTs and in dense urban cores a BG can be as small as a few city blocks. Therefore, BG-level data likely leads to more accurate measures of income in treated and control areas than CT-level data.[11] The cost of using BG-level data to generate *M* instead of CT-level data is a loss in temporal breadth in the analyses. When using BT-level income data I had to set $\underline{t} = 1990$ because nation-wide digital maps of BGs only start with the 1990 US Census. Conversely, nation-wide maps of CTs begin with the 1980 US Census, thereby allowing me to set $\underline{t} = 1980$ when I used CT-level data. This also means my pre-trend analyses largely relied on CT-level data given it provides multiple pre-treatment estimates of (4.1) while the BG-level data does not.

Until the mid-2000s the US Census Bureau only measured income at the BG and CT levels during a decadal census. However, starting in 2005 the Bureau estimated income at these spatial levels every year with the ACS. However, because the ACS sample sizes are relatively small, the Bureau recommends using several years of ACS data to measure income in a BG or CT at any point in time. Therefore, the index *t* in model (4.1) post-2005 does not include every year but rather 2010, based on data collected between 2006-2010, and 2017, based on data collected between 2013-2017 (i.e., $\bar{t}$ = 2017).

I estimated (4.1) for the urban area (u) – treatment year (τ) combinations of LA-2003; LA-2005; Phoenix-2008; Denver-1994; Denver-2006, and Minneapolis-2004. All treatments involve the opening of new light rail stations save LA-2005, which is centered on the opening of

[11] For example, it is not uncommon for a treated or control area to be contained within one CT and for the CT area outside the treated or control polygon to be substantial. In this case, the income measure for the treated or control polygon could be heavily influenced by income measures from outside the polygon. However, this same polygon is likely to better follow the contours of a set of nested BGs, meaning the BG-level income measure for the polygon will be less influenced by income data from outside the polygon than the CT-level income measure for the polygon.

rapid bus line stations. For the LA (MSA category 1), Denver (MSA category 2), and Minneapolis (MSA category 2) events I create a control set of nonoverlapping polygons centered on random points along historic streetcar and trolley lines (Severen 2018, CCDPWPP 2019, Metropolitan Council 2007). Unfortunately, I cannot find a map of historic Phoenix (MSA category 2) streetcar lines that ended service in 1948. In addition, I found maps of planned post-2017 light rail station expansions in Phoenix, Denver, and Minneapolis. In Phoenix post-2017 stations were slated to open in 2019, 2021, 2023, and 2030. In Denver post-2017 stations were slated to open in 2019 and 2020. In Minneapolis post-2017 stations were slated to open in 2023 (see SI Text 3 for instructions on replicating my results).

4b. Interpreting event-study model estimates

Before estimating (4.1) with ordinary least squares I needed to drop the dummy variable $\mathbf{1}[Treat]_c$ because it equals the sum of the treated areas' fixed effects (Clay et al. (2016) also drop their version of $\mathbf{1}[Treat]_c$ before estimating their version of model (4.1)). I also dropped the D_{ct}^{τ} associated with the period immediately prior to treatment, indicated by π, to avoid perfect multicollinearity. In all event-study analyses I conduct save one, year π is 2000 (i.e., $2000 < \tau < 2010$). In these cases, $\hat{\delta}_{1980}$ and $\hat{\delta}_{1990}$ measure the relative difference in average income or poverty rate between treated and control areas τ – 1980 and τ – 1990 years, respectively, before station openings, and $\hat{\delta}_{2010}$ and $\hat{\delta}_{2017}$ measure the relative difference in average income or poverty rate between treated and control areas 2010 – τ and 2017 – τ years, respectively, after station openings ($\hat{\delta}_{1980}$'s are not produced when I use BG-level income data). Equivalent $\hat{\delta}_{1980}$ and $\hat{\delta}_{1990}$ would suggest a common pre-trend between treated and control areas. Further, $\hat{\delta}_{2010}$ and $\hat{\delta}_{2017}$ that are significantly different from $\hat{\delta}_{1980}$ and $\hat{\delta}_{1990}$, both in magnitude and statistically, would suggest that the opening of light rail stations *caused* a change in the treated acres' average income dynamics relative to contemporaneous income dynamics in the control areas.

Denver-1994 is the one aforementioned exception to $2000 < \tau < 2010$. In this case I drop the D_{ct}^{τ} associated with 2000, the period immediately *after* treatment. Despite this change, $\hat{\delta}_{2000}$ is still set equal to 0, $\hat{\delta}_{1980}$ and $\hat{\delta}_{1990}$ still map out pre-trends, and $\hat{\delta}_{2010}$ and $\hat{\delta}_{2017}$ still indicate average (relative) treatment effects.

4c. The no anticipation condition

Well-identified event-studies also satisfy the 'no anticipation condition' (Abraham and Liyang 2018). For example, assume *u*'s political leaders announced in 1998 that a new light rail station would open in 2005 at point *x* in *u*. Suppose *u* residents and developers immediately began to make housing location and home building decisions based on this intention. In this case I could not credibly claim that $\hat{\delta}_{2000}$ was a pre-treatment measure as it would identify some of the anticipated impact of the station opening. However, there are several reasons to suspect that the 'no anticipation condition' is not violated in my event-studies. First, the gaps between the year π and τ in my analyses are 3, 4, 5, 6, and 8 years. Therefore, if anticipatory behavior begins a year, two, or three before station opening it would not be soon enough to affect the last pre-treatment coefficient. For example, Hurst and West (2014) find faster than expected land use change in the Twin Cities Blue Light Rail Line corridor (Minneapolis-2004) only began to occur in 2001, a year *after* line construction began in 2000 (i.e., $\hat{\delta}_{2000}$ in this case is not affected by anticipatory behavior). Further, Cao and Porter-Nelson (2016) find that building permits did not increase around a new light rail line, the Twin Cities Green line (not studied here), which opened in 2014, until a full funding agreement from the federal government was issued in 2011 (a three-year gap).

Second, there can be a multi-year lag between the decision to build new housing and the execution of that the decision in US urban areas, especially in heavily regulated inner-city cores that host most light rail stations (Bahadir and Mykhaylova 2014, Gyourko et al. 2008). For example, while developers may immediately have begun planning for a 2005 station opening when announced in 1998, their choices may not have affected HH relocation decisions until new residential units were built in 2001 or later. In other words, if any new station in my analysis was announced in the late 1990s, any changes on the landscape that HHs could react to were not likely until after 2000.

Third, the locations of some recently built light rail stations were not finalized by regulators until after 2000. For example, the Metropolitan Council, a regional transit authority in the Twin Cities area, did not approve Minneapolis-2004 station area plans until 2001 despite a 2004 opening date (Newberg 2004). Phoenix voters approved an increase in the city's sales tax to fund its 2008 light rail line in 2000; serious consideration of station locations only began after this vote (Arizona Rail Passenger Association 2007). The path of Denver's E line, which opened in 2006 (Denver-2006), was only approved by voters in late 1999 (Schneider 2005). Denver-based consumer and business reaction to this news was likely several years in the making, well beyond the 2000 census.

I also assume the no anticipation condition in the control areas centered on post-2017 light rail stations for similar reasons. I assume that people's housing location decisions and housing and amenity supply and location across *u* as of 2017 had not yet incorporated station openings that, at the earliest, were not to open for another 2 years. This is not to say that planning in city and development firm offices given this information had not yet begun by 2017 in Denver, Minneapolis, and Phoenix. Rather, I assume this planning have not yet shaped the actual decisions of the cities' residents nor affected the realized supply and location of housing and amenities in these cities as of 2017.

4d. Event-study results

Overall, estimates of (4.1) across the various urban area, treatment year combinations suggest that areas immediately surrounding light rail or rapid bus line stations that opened at the beginning of the 21st century did not become magnets for poor HHs relative to the control areas in the same urban area.

4d.1. Results with historic streetcar and trolley controls

In LA, areas treated with new PT stations in 2003 (Table 8) and 2005 (Table 9) experienced significant increases in average median HH income relative to the randomly selected control polygons centered on LA's historic streetcar network. Further, given there was no statistical difference in CT-level average median HH income between the two sets of polygons before treatment (i.e., similar pre-trends), there is some evidence to suggest that treatment caused this increase in nearby median HH income. Conversely, treatment seems to have had little impact on the relative differences in per capita income and poverty rates across the two sets of areas. While treated LA areas experienced an increase in per capita income relative to the control areas along the historic PT network, the treatment effect was not statistically different from zero. Finally, the poverty rate gap between treated and control areas either fell (improved in the treated areas) after treatment (τ = 2003) or did not change (τ = 2005).

In Minneapolis, areas treated with new PT stations in 2004 (Table 10) experienced no statistically significant change in average median HH income or per capita income relative to the randomly selected control polygons along its historic trolley network. Generally, average differences in CT-level income between these two sets of areas before treatment were not statistically different from 0 either. However, these is some evidence to suggest that treatment reduced poverty rates relative to rates in the historic trolley network control areas. Before treatment, average poverty rates were relatively higher in the treated areas relative to the control areas. By 2017, the statistical difference in average poverty rates between the two types of areas was 0. However, given that the trends in relative average poverty rates at the CT-level were not common prior to treatment, poverty was relatively increasing in treated areas prior to treatment, I cannot make an argument for causality in the poverty rate case.

In the Denver-1994 case study income and poverty rates did not differ relative to trends in the controls centered on the city's historic streetcar network (Table 11). However, in Denver's 2006 event-study analysis with the historic PT network controls I find the one instance where PT treatment unequivocally led to lower incomes in the affected neighborhoods relative to trends in the control neighborhoods (Table 12). Whether measured with median HH income or per capita income, areas around stations that opened in 2006 experienced a large and statistically significant decrease in average income post-treatment compared to the control areas (before treatment, there was generally no statistical difference in CT-level average incomes between the two sets of areas). Poverty rates relatively increased in these treated areas as well post-treatment.

4d.2. Results with post-2017 controls

In Denver's 1994 event-study analysis with future station area controls (Table 11), treatment in 1994 created significant increases in average incomes and reductions in poverty rates relative to trends in control areas. The effect of treatment on poverty rates is particularly striking. On average, before treatment, the CT-level poverty rates in treated areas was approximately 10 percentage points higher than in control areas. Twenty-three years after treatment, the average poverty rate in treated areas was approximately 10 percentage points *lower* than in the control areas.

In Denver's other event-study analysis with future station area controls (Table 12), treatment has no consistent impact on treated area average incomes and poverty rates relative to the control averages. Similarly, in Minneapolis' event-study analysis with post-2017 controls (Table 10), treatment has no consistent impact on relative average income levels and poverty rates.

In Phoenix's event-study analysis with post-2017 station area controls (Table 13), I find that treatment in 2008 generated a statistically significant increase in surrounding area average median HH and per capita income as of 2017 relative to the areas that will receive treatment in the future. Before treatment the CT-level relative difference in incomes was either statistically 0 or slightly lower in treated areas (e.g., weak evidence of common pre-trends). Treatment in this case either had no effect on relative average poverty rates or slightly decreased poverty in affected neighborhoods relative to the control. Before treatment there was no statistical difference in average CT-level poverty rates across the two sets of polygons.

5d.3. Results with treatment area buffer controls

In my analyses with treated area buffer controls I test whether areas right next to a station trend differently than areas *ever-so-slightly* further away. In every one of the event-study analyses I found the differences in average income and poverty rate across the treated and control areas prior to treatment were statistically equivalent to 0. This is not surprising given that a treated area and its buffer likely cover the same neighborhoods. Further, other the Phoenix-2008 case (Table 13), I find no evidence that treated areas experienced income or poverty trends different than their buffer areas post-treatment. In Phoenix there is some evidence to suggest the 2008 opening of light rail stations caused areas right next to the stations to become richer throughout the 2010s than the areas *ever-so-slightly* further away.

5.e. Summarizing third analysis results

Unlike Glaser et al. (2008)'s conclusion that CTs treated with new transit lines subsequently became more impoverished, I find little evidence to suggest that areas treated with new PT stations in the 2000s subsequently became more impoverished. I have already suggested that Glaeser et al. (2008)'s poor identification strategy may be one reason for these contradictory results.

Of course, I could be wrong and the differences in results could be explained by other phenomena. First, the event-study windows do not align. A trend of poor HHs concentrating around PT infrastructure could have largely ended by 2000, meaning Glaeser et al. (2008)'s analysis would show such a trend and mine would not. Second, Glaser et al. (2008) includes areas with new transit lines, not just a new station, in their treated set. It is possible that if I had done the same then I would have generated results more in line with theirs. Third, the cities I use in my analysis are not the same as the 16 that Glaser et al. (2008) use in their event-study analysis. Fourth, my analysis focuses on new light rail and rapid-bus stations. My results might be more in line with Glaser et al. (2008) if I had included conventional bus line expansions in my analysis. For example, Pathak et al. (2017) find that areas in the Atlanta metropolitan area treated with conventional bus PT became slightly poorer relative to untreated control areas.

6. Discussion and conclusions

I have found Glaeser et al. (2008)'s claim that "[p]ublic transportation usage appears to strongly predict poverty and to explain a substantial amount of the connection between proximity and poverty" to be unsubstantiated by 2017 data.

My estimates of commute speeds by technology and by HH income class suggest that observed patterns of HH income ring sorting in most US urban areas cannot be explained by a monocentric city model solely animated by commute technology. Instead, the empirical evidence as of 2017 suggests that the spatial ordering of similar HH income clumps in US cities is largely explained by phenomenon other than poor commuters relying on PT. Of course, my results do not rule out the possibility that poorer HHs tend to cluster near PT networks for reasons other than easy access to a preferred commute technology. For example, non-working members of a poor HH may find PT their preferred method for accessing urban amenities and educational opportunities. Or the poor commuter may prefer to use PT for non-work travel and therefore want to locate their HH near PT networks.

However, my event-study analyses of recently opened light rail and rapid-bus stations do not support this alternative hypothesis. Unlike Glaeser et al. (2008), I do not find that recent additions to urban PT networks cause their surrounding neighborhoods to become poorer over time relative to neighborhoods not treated by additional PT infrastructure. At least in the cities that I investigated and the PT technologies I considered, poorer HHs do not move to new PT infrastructure more than middle class and rich HHs.

Finally, Glaeser et al. (2008)'s conclusion that a substantial increase in income at poor HHs should be associated with a massive shift from public transportation to driving is not borne out when a mode choice model is estimated with 2017 data. In general, American commuters are not likely to switch commuting modes, at least in the short run, no matter the exogenous shock. Most commute by car and will continue to commute by car. In the densest neighborhoods of its largest cities, the one type of place in the US where some mode switching among poor commuters does occur, the impetus for mode choice change cannot simply be attributed to changes in HH income.

I finished this analysis just as the virus that causes COVID-19 swept across the world. Many are predicting that HH location decisions and commuting behavior across US urban areas will permanently change due to the pandemic. In other words, some the trends and patterns that I discuss in this paper may not apply for the next few years, if ever again. I identify three of the COVID-19-related shocks that could change HH location decision-making and commuter behavior across US urban areas for years to come.

White-collar telecommuting is projected to become more widespread from here on out (Guyot and Sawhill 2020) for several reasons. First, many US workers and their employers are discovering that this mode protects worker health, due to social distancing measures, while maintaining and even increasing worker productivity and satisfaction (Dutcher 2012, Bloom et al. 2015, Davis and Green 2020). Second, the emergence of several positive externalities associated with more telecommuting, including reductions in health-damaging criteria pollutant and climate change-causing greenhouse gases emissions (Cadotte 2020, Cicala et al. 2020) and less congestion (Tomer and Fishbane 2020), has led to some nascent policy proposals to mandate or incentivize higher rates of telecommuting even when the crisis has subsided (Nguyen 2020). The rising popularity of the telecommuting mode is likely to reduce commuting by car and PT in the US for many years and reduce the role of workplace location in HH location decision-making.

Second, the pandemic is likely to make PT commuting even less popular in the future above and beyond the displacement caused by the rise in telecommuting. First, upcoming pandemic-generated fiscal crises are likely to lead national and regional-level governments to reduce investments in PT. Second, beliefs that PT systems are strong vectors of disease spread will dissuade some commuters from using these system for many years to come (Hawkins

2020).[12] These severe reductions in PT investment and PT revenue is likely to lead to even more service reductions and accelerated infrastructure depreciation (Walker 2020), steadily making PT a less preferred commute option for more and more people around the world.

Third, the drop in PT during the COVID 19 pandemic has coincided with an increase in rates of bicycling and walking (Schwedhelm et al. 2020). Some cities, such as London and Seattle, are responding by opening more bike and walking lanes at the expense of road capacity (Department for Transport and The Rt Hon Grant Shapps MP 2020, Jackson 2020). It is uncertain how much of this short-term trend in greater cycling and more cycling infrastructure is fulfilling recreational versus commuting demand during lockdown orders. However, past research has shown that as more people experience biking in cities and as cities expand bike lanes, commuting by bicycle increases (Hirsch et al. 2017, Nelson and Allen 1997). Therefore, lingering trepidation over using PT systems post-crisis and more experience with safe biking during the immediate crisis could lead to a permanent increase in commuting by bicycle into the future.

In conclusion, I have argued that 1) poor commuters are not as reliant on PT and 2) the spatial allocation of PT systems in US urban areas does less to explain the spatial pattern of HH incomes within these urban areas as Glaser et al. (2008) would have us believe. I suspect that the COVID-19 pandemic will make PT even less relevant to American commuters in the future,

[12] Whether PT systems are effective spreaders of COVID-19 is debatable. Harris (2020) claims that the New York City subway system did much to spread the disease in the city. However, Almagro and Orane-Hutchinson (2020) "…show that after controlling for occupations, length of commute and the use of public transport are not significant [determinants of the fraction of tests showing a positive result across NYC zip codes]" (p. 2). (Also see Furth 2020). Further, cities like Seoul, Taipei, and Singapore have avoided large spread of the disease despite the continued operation of their subway systems (Park 2020).

thereby further reducing any links between PT and HH income spatial patterns that exist in US urban areas.

7. Acknowledgments

Thank you to Christopher Severen for sharing his digital maps of historical streetcar networks in Los Angeles.

8. References

Abraham S., Sun L., 2018, April 16. Estimating Dynamic Treatment Effects in Event Studies With Heterogeneous Treatment Effects. http://dx.doi.org/10.2139/ssrn.3158747.

American Community Survey, U.S. Census Bureau (ACS), 2017. 2013-2017 5-year Estimates, Table S1903, Median Income in the Past 12 Months (In 2017 Inflation-Adjusted Dollars). Generated by Erik Nelson 21 March 2019. https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2017/5-year.html

Almagro M., Orane Hutchinson A., 2020, April 10. The Determinants of the Differential Exposure to COVID-19 in New York City and Their Evolution Over Time. http://dx.doi.org/10.2139/ssrn.3573619.

Arizona Rail Passenger Association, 2007. A Brief History of Transportation Elections in the Valley of the Sun (1960-2000). The Information Group, Phoenix. https://web.archive.org/web/20070812211140/http://www.azrail.org/trains/transit/transit-elections/. Accessed May 14, 2020.

Bahadir B., Mykhaylova O., 2014. Housing market dynamics with delays in the construction sector. Journal of Housing Economics 26, 94-108. https://doi.org/10.1016/j.jhe.2014.09.005.

Bhat C.R., Sardesai R., 2006. The impact of stop-making and travel time reliability on commute mode choice. Transportation Research Part B: Methodological 40 (9), 709-730. https://doi.org/10.1016/j.trb.2005.09.008.

Bhatta B.P., Larsen O.I., 2011. Errors in variables in multinomial choice modeling: A simulation study applied to a multinomial logit model of travel mode choice. Transport Policy 18 (2), 326-335. https://doi.org/10.1016/j.tranpol.2010.10.002.

Bloom N., Liang J., Roberts J., Ying Z.J., 2015. Does Working from Home Work? Evidence from a Chinese Experiment. The Quarterly Journal of Economics, 130 (1), 165–218. https://doi.org/10.1093/qje/qju032.

Brooks L., Lutz B., 2019. Vestiges of Transit: Urban Persistence at a Microscale. The Review of Economics and Statistics, 101 (3), 385-399. https://doi.org/10.1162/rest_a_00817.

Cadotte M., 2020, March 30. Early evidence that COVID-19 government policies reduce urban air pollution. https://doi.org/10.31223/osf.io/nhgj3.

Cartenì A., Cascetta E., de Luca S., 2016. A random utility model for park & carsharing services and the pure preference for electric vehicles. Transport Policy 48, 49-59. https://doi.org/10.1016/j.tranpol.2016.02.012.

Cicala S., Holland S.P., Mansur E.T., Muller N.Z., Yates A.J., 2020, May. Expected Health Effects of Reduced Air Pollution from COVID-19 Social Distancing. National Bureau of Economic Research, NBER Working Paper No. 27135, https://doi.org/10.3386/w27135.

Cao X., Porter-Nelson D., 2016. Real estate development in anticipation of the Green Line light rail transit in St. Paul. Transport Policy 51, 24-32. https://doi.org/10.1016/j.tranpol.2016.01.007.

City and County of Denver, Public Works - Policy and Planning, 2019, December 17. Abandoned Trolley Tracks. City and County of Denver, Technology Services / DenverGIS Data, v. 1.0.4. https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-abandoned-trolley-tracks. Accessed May 18, 2020.

Clay K., Lewis J., Severnini E., 2016, April. Canary in a Coal Mine: Infant Mortality, Property Values, and Tradeoffs Associated with Mid-20th Century Air Pollution. National Bureau of Economic Research, NBER Working Paper No. 22155, https://doi.org/10.3386/w22155.

Davis M.F., Green J., 2020, April 23. Three Hours Longer, the Pandemic Workday Has Obliterated Work-Life Balance. Bloomberg News Service, https://www.bloomberg.com/news/articles/2020-04-23/working-from-home-in-covid-era-means-three-more-hours-on-the-job. Accessed May 18, 2020.

Department for Transport and The Rt Hon Grant Shapps MP, 2020, May 9. £2 billion package to create new era for cycling and walking. GOV.UK. https://www.gov.uk/government/news/2-billion-package-to-create-new-era-for-cycling-and-walking. Accessed May 18, 2020.

Dutcher E.G., 2012. The effects of telecommuting on productivity: An experimental examination. The role of dull and creative tasks. Journal of Economic Behavior & Organization 84 (1), 355-363, https://doi.org/10.1016/j.jebo.2012.04.009.

El-Assi W., Salah Mahmoud M., Nurul Habib K., 2017. Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. Transportation 44, 589-613. https://doi.org/10.1007/s11116-015-9669-z.

Eluru N., Chakour V., El-Geneidy A.M., 2012. Travel mode choice and transit route choice behavior in Montreal: insights from McGill University members commute patterns. Public Transport 4, 129-149, https://doi.org/10.1007/s12469-012-0056-2.

Federal Highway Administration (FHA), 2017. 2017 National Household Travel Survey, U.S. Department of Transportation, Washington, DC. https://nhts.ornl.gov.

Frank L., Bradley M., Kavage S., Chapman J., Lawton T.K., 2008. Urban form, travel time, and cost relationships with tour complexity and mode choice. Transportation 35, 37-54. https://doi.org/10.1007/s11116-007-9136-6.

Frank L., Pivo G.,1994. Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, and walking. Transportation Research Record 1466, 44-52.

Furth S., 2020, April 19. Automobiles seeded the massive coronavirus epidemic in New York City. Market Urbanism. https://marketurbanism.com/2020/04/19/automobiles-seeded-the-massive-coronavirus-epidemic-in-new-york-city/. Accessed May 19, 2020.

Glaeser E.L., Kahn M.E., Rappaport J., 2008. Why do the poor live in cities? The role of public transportation. Journal of Urban Economics 63 (1), 1-24. https://doi.org/10.1016/j.jue.2006.12.004.

Glazier R.H., Creatore M.I., Weyman J.T., Fazli G., Matheson F.I., Gozdyra P., Moineddin R., Kaufman-Shriqui V., Booth G.L., 2014. Density, destinations or both? A comparison of measures of walkability in relation to transportation behaviors, obesity and diabetes in Toronto, Canada. PLoS ONE 9 (1), e85295. https://doi.org/10.1371/journal.pone.0085295.

Goodman-Bacon A., 2019. Difference-in-differences with variation in treatment timing. https://my.vanderbilt.edu/ajgb/research-in-progress/. Accessed May 19, 2020.

Guyot K., Sawhill I.V., 2020, April 6. Telecommuting will likely continue long after the pandemic. Brookings Institution. https://www.brookings.edu/blog/up-front/2020/04/06/telecommuting-will-likely-continue-long-after-the-pandemic/. Accessed May 19, 2020.

Gyourko J., Saiz A., Summers A., 2008. A new measure of the local regulatory environment for housing markets: the Wharton residential land use regulatory index. Urban Studies 45 (3), 693-729. https://doi.org/10.1177/0042098007087341.

Harris J.E., 2020, April. The Subways Seeded the Massive Coronavirus Epidemic in New York City. National Bureau of Economic Research, NBER Working Paper No. 27021. https://doi.org/10.3386/w27021.

Hawkins A.J., 2020, March 13. Coronavirus is taking a big bite out of public transportation ridership in the US. The Verge. https://www.theverge.com/2020/3/13/21179032/public-transportation-coronavirus-covid19-ridership-nyc-sf-la-dc. Accessed May 19, 2020.

Hirsch J.A., Meyer K.A., Peterson M., Zhang L., Rodriguez D.A., Gordon-Larsen P., 2017. Municipal investment in off-road trails and changes in bicycle commuting in Minneapolis, Minnesota over 10 years: a longitudinal repeated cross-sectional study. International Journal of Behavioral Nutrition and Physical Activity 14, 21. https://doi.org/10.1186/s12966-017-0475-1.

Holian M.J., Kahn M.E., 2015. Household carbon emissions from driving and center city quality of life. Ecological Economics 116, 362-368. https://doi.org/10.1016/j.ecolecon.2015.05.012.

Hurst N.B., West S.E., 2014. Public transit and urban redevelopment: The effect of light rail transit on land use in Minneapolis, Minnesota. Regional Science and Urban Economics 46, 57-72. https://doi.org/10.1016/j.regsciurbeco.2014.02.002.

Jackson A., 2020, May 7. Seattle to permanently close 20 miles of streets to traffic so residents can exercise and bike on them. CNN. https://www.cnn.com/travel/article/seattle-streets-closed-stay-healthy-trnd/index.html. Accessed May 19, 2020.

Kelker, De Leuw & Company, 1925. Report and Recommendations on a Comprehensive Rapid Transit Plan for the City and County of Los Angeles. http://libraryarchives.metro.net/dpgtl/trafficplans/1925_comprehensive_rapid_transit_plan_los_angeles.pdf. Accessed May 19, 2020.

Kenworthy J.R., Laube F.B., Newman P., 1999. An International Sourcebook of Automobile Dependence in Cities, 1960-1990. University Press of Colorado, Boulder, CO.

LeRoy S.F., Sonstelie, J., 1983. Paradise lost and regained: Transportation innovation, income, and residential location. Journal of Urban Economics, 13(1), 67-89.

Limtanakool N., Dijst M., Schwanen M., 2006. The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium- and longer-distance trips. Journal of Transport Geography 14(5), 327-341. https://doi.org/10.1016/j.jtrangeo.2005.06.004.

Manson S., Schroeder J., Van Riper D., Ruggles S., 2019 IPUMS National Historical Geographic Information System: Version 14.0, Minneapolis, MN, IPUMS. http://doi.org/10.18128/D050.V14.0.

McFadden D., 1977. Quantitative Methods for Analyzing Travel Behaviour of Individuals: Some Recent Developments," Cowles Foundation Discussion Papers 474, Cowles Foundation for Research in Economics, Yale University. https://ideas.repec.org/p/cwl/cwldpp/474.html. Accessed May 20, 2020.

Metropolitan Council, 2007, November 1. Historical Transit Routes - Twin Cities Metropolitan Area. Minnesota Geospatial Commons. https://gisdata.mn.gov/dataset/us-mn-state-metc-trans-historical-transit-routes.

Nelson A.C., Allen D., 1997. If You Build Them, Commuters Will Use Them: Association Between Bicycle Facilities and Bicycle Commuting. Transportation Research Record 1578 (1), 79-83. https://doi.org/10.3141/1578-10.

Newberg S., 2004, May. Light Rail Comes to Minnesota. Commuters get set to greet the 12-mile Hiawatha Line. American Planning Association. https://web.archive.org/web/20111011062726/http://community-wealth.org/_pdfs/articles-publications/tod/article-newberg.pdf. Accessed May 21, 2020.

Newman P., Kenworthy J., 1989. Gasoline Consumption and Cities: A Comparison of U.S. Cities with a Global Survey. Journal of the American Planning Association 55 (1), 24-37. https://doi.org/10.1080/01944368908975398.

Nguyen C., 2020, May 4. Coronavirus Impact: Santa Clara Co. proposal would allow more employees to work from home after pandemic. ABC 7 News. https://abc7news.com/working-remotely-from-home-tips-work-jobs-santa-clara-county/6151209/. Accessed May 21, 2020.

Park J., 2020. Changes in Subway Ridership in Response to COVID-19 in Seoul, South Korea: Implications for Social Distancing. Cureus 12 (4), e7668. https://doi.org/10.7759/cureus.7668.

Pathak R., Wyczalkowski C.K., Huang X., 2017. Public transit access and the changing spatial distribution of poverty. Regional Science and Urban Economics 66, 198-212. https://doi.org/10.1016/j.regsciurbeco.2017.07.002.

Raghunathan T. E., Lepkowski J. M., Hoewyk J. Van, Solenberger, P., 2001. A multivariate technique for multiply imputing missing values using a sequence of regression models. Survey Methodology 27, 85-95

Scheiner J., Holz-Rau, C., 2012. Gendered travel mode choice: a focus on car deficient households. Journal of Transport Geography 24, 250-261. https://doi.org/10.1016/j.jtrangeo.2012.02.011.

Schneider K., 2005, June 16. In Denver's Transit Breakthrough, A Lesson For Detroit. Great Lakes Bulletin News Service, Michigan Land Use Institute. http://www.mlui.org/mlui/news-views/articles-from-1995-to-2012.html?archive_id=544#.Xsa1NsB7k2w. Accessed May 21, 2020.

Schoner J.E., Levinson D.M., 2014. The missing link: Bicycle infrastructure networks and ridership in 74 US cities. Transportation 41, 1187–1204. https://doi.org/10.1007/s11116-014-9538-1.

Schwedhelm A., Li W., Harms L., Adriazola-Steil C., 2020, April 17. Biking Provides a Critical Lifeline During the Coronavirus Crisis. World Resources Institute. https://www.wri.org/blog/2020/04/coronavirus-biking-critical-in-cities. Accessed May 21, 2020.

Severen C., 2018. Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification. Working Papers 18-14, Federal Reserve Bank of Philadelphia, revised March 2019. https://doi.org/10.21799/frbp.wp.2018.14.

Shen Q., Chen P., Pan H., 2016. Factors affecting car ownership and mode choice in rail transit-supported suburbs of a large Chinese city. Transportation Research Part A: Policy and Practice 94, 31-44. https://doi.org/10.1016/j.tra.2016.08.027.

Tomer A., Fishbane L., 2020, May 1. Coronavirus has shown us a world without traffic. Can we sustain it? Brookings Institution. https://www.brookings.edu/research/coronavirus-has-shown-us-a-world-without-traffic-can-we-sustain-it/. Accessed May 21, 2020.

Walker J., 2020, March 10. Covid-19: What if Transit Runs Out of Money? Human Transit [blog]. https://humantransit.org/2020/03/covid-19-what-if-transit-runs-out-of-money.html. Accessed May 21, 2020.

Whalen K.E., Páez A., Carrasco J.A., 2013. Mode choice of university students commuting to school and the role of active travel. Journal of Transport Geography 31, 132-142. https://doi.org/10.1016/j.jtrangeo.2013.06.008.

Wing C., Simon K., Bello-Gomez R.A., 2018. Designing difference in difference studies: best practices for public health policy research. Annual Review of Public Health 39, 453-469. https://doi.org/10.1146/annurev-publhealth-040617-013507.

Figures

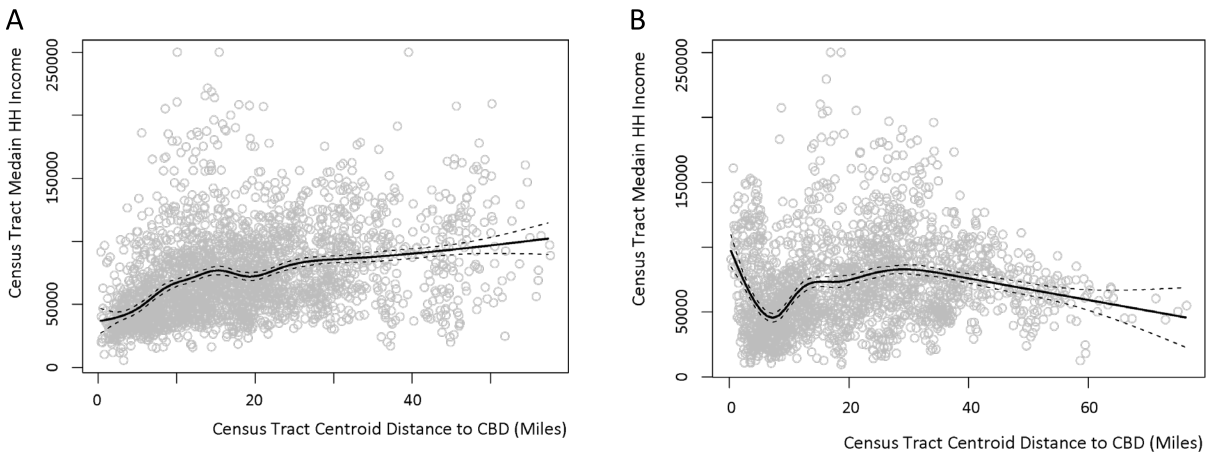


Figure 1. Census tract-level median household income splines for the (A) Los Angeles and (B) Chicago core based statistical areas. Each plot point indicates a census tract's median HH income circa 2017 (ACS 2017) and its Euclidean distance (in miles) to the CBSA's central business district (CBD) (Holian and Kahn 2015). The splines are natural splines with 8 degrees of freedom (the dotted lines give the 95% confidence level bounds). 2017 census tract location data comes from the US Census' 2017 Gazetteer Files. Run the R script *Figure1* to replicate these figures.

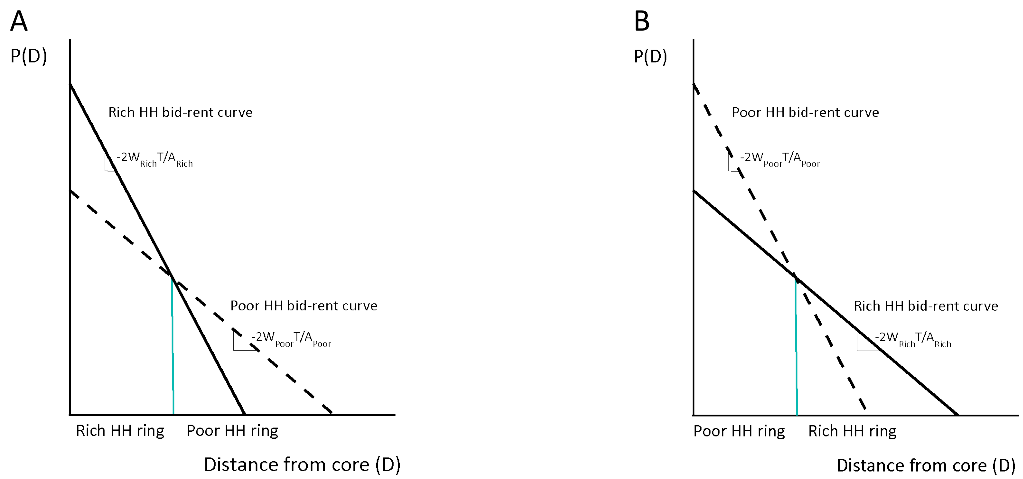


Figure 2. Sorting of households in theoretical urban areas. (A) In this urban area the rich HHs' bid-rent curve is steeper ($\varepsilon_Y^W > \varepsilon_Y^A$), and if the urban area is to have both poor and rich HHs, then the rich settle in the inner ring and the poor in the outer ring. (The pattern in Chicago would seem to suggest four rings: richer HHs, then poorer HHs, then more well-to-do HHs, and finally poorer HHs again. My application of the theory is limited to the first two rings. However, the theory could be extended to include additional bud-rent curves for other classes HHs.) (B) In this urban area the poor HHs bid-rent curve is steeper ($\varepsilon_Y^W < \varepsilon_Y^A$), and if the urban area has both poor and rich HHs, then the poor settle in the inner ring and the rich in the outer ring.

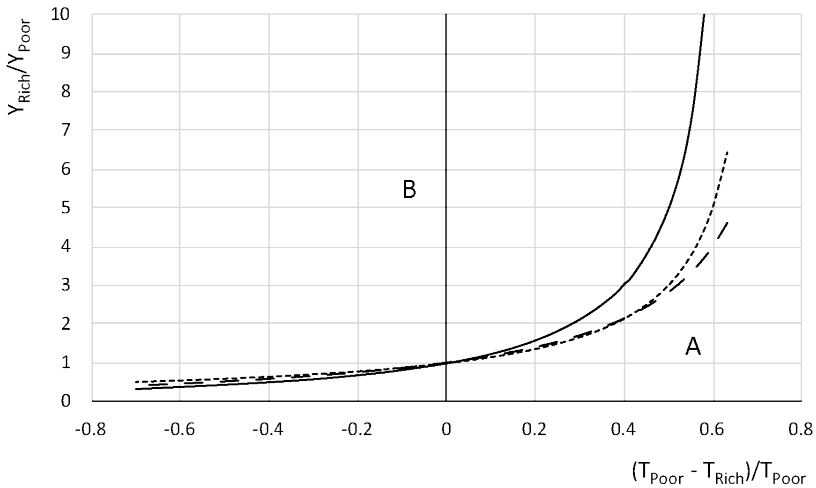


Figure 3. For HHs in the monocentric city's first ring to be poor (i.e., $\varepsilon_Y^A + \frac{T_{Poor}-T_{Rich}}{T_{Poor}}\left(\frac{Y_{Poor}}{Y_{Rich}-Y_{Poor}} + \varepsilon_Y^W\right) > \varepsilon_Y^W$), the combination of $\frac{T_{Poor}-T_{Rich}}{T_{Poor}}$ and Y_{Rich}/Y_{Poor} has to be in space A. For HHs in the monocentric city's first ring to be rich (i.e., $\varepsilon_Y^A + \frac{T_{Poor}-T_{Rich}}{T_{Poor}}\left(\frac{Y_{Poor}}{Y_{Rich}-Y_{Poor}} + \varepsilon_Y^W\right) < \varepsilon_Y^W$), the combination of $\frac{T_{Poor}-T_{Rich}}{T_{Poor}}$ and Y_{Rich}/Y_{Poor} has to be in space B. In the graph $\varepsilon_Y^W = 0.75$ and $\varepsilon_Y^A = 0.25$ (solid line); $\varepsilon_Y^W = 0.75$ and $\varepsilon_Y^A = 0.25$ (dotted line); and $\varepsilon_Y^W = 1$ and $\varepsilon_Y^A = 0.25$ (dashed line). See appendix for details.

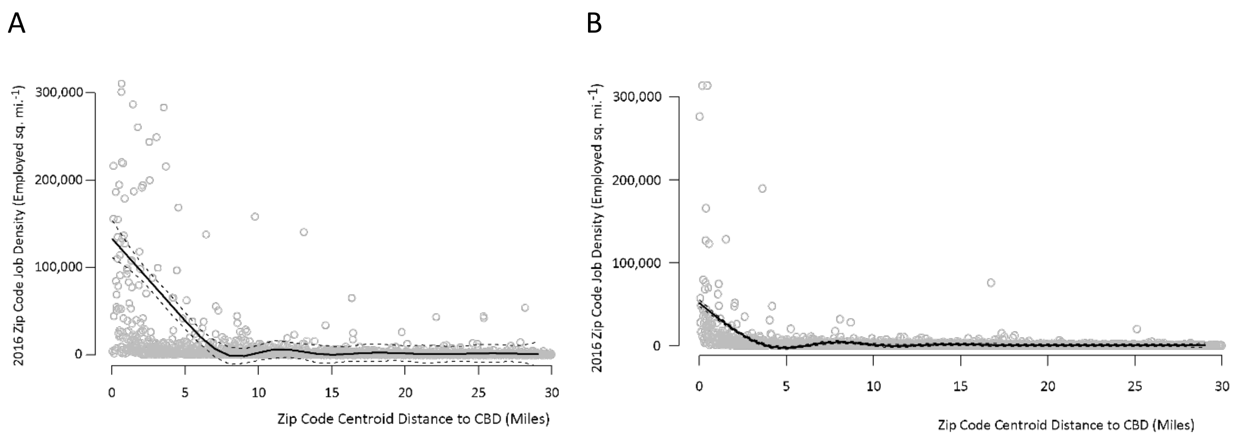


Figure 4: **Employment density gradients at the zip code level as a function of zip code distance from the CBD for all urban areas in MSA categories 1 (A) and 2 (B).** Estimated and 5th and 95th percentile splines are "natural" splines with 8 degrees of freedom. Run the R scripts *Figure4A* and *Figure4B* to replicate these figures.

A B

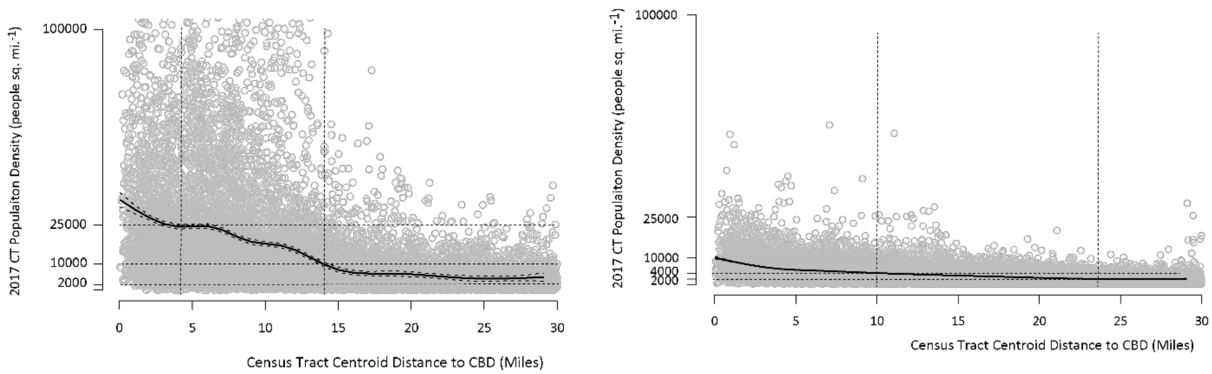


Figure 5: **Population density gradients at the CT level as a function of CT distance from the CBD for all urban areas in MSA categories 1 (A) and 2 (B).** Estimated and 5th and 95th percentile splines are "natural" splines with 8 degrees of freedom. The dotted lines indicate where a spline intersects a population density threshold used in the 2017 NHTS. Therefore, the first 5 mile interval and the 5 to 15-mile interval from the CBD form the first and second urban rings, respectively, in the collection of MSA category 1 urban areas. Further, the first 10 mile interval and the 10 to 23-mile interval from the CBD form the first and second urban rings, respectively, in the collection of MSA category 2 urban areas. Run the R scripts *Figure5A* and *Figure5B* to replicate these figures.

A B

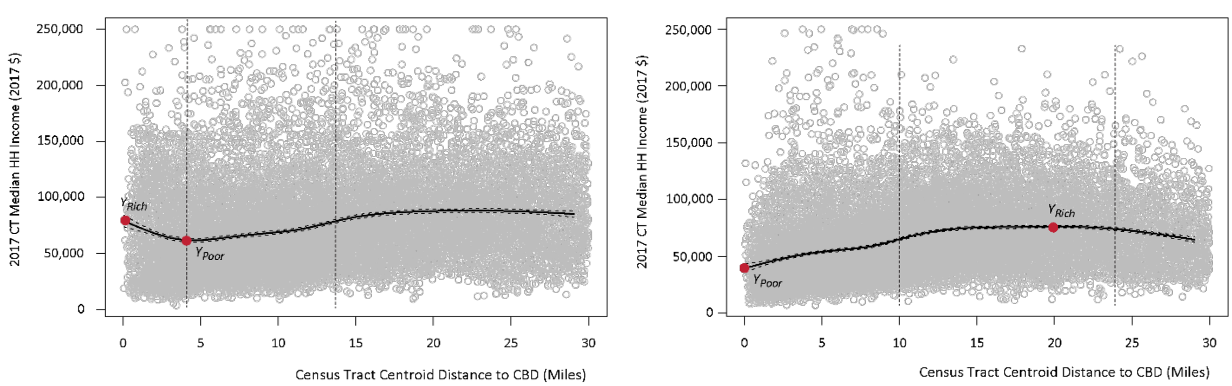


Figure 6: **HH income gradients at the CT level as a function of CT distance from the CBD for all urban areas in MSA categories 1 (A) and 2 (B).** Estimated and 5th and 95th percentile splines are "natural" splines with 8 degrees of freedom. The dotted lines indicate where a spline intersects a population density threshold used in the 2017 NHTS. In addition, I set Y_{Poor} and Y_{Rich} for the collection of urban areas in each MSA category equal to the lowest and highest points, respectively, of that category's median HH income spline (within the first two urban rings). Y_{Poor} = 61,681 (mile 4) and Y_{Rich} = 78,968 (mile 0) in MSA 1 category urban areas. Further, Y_{Poor} = 39,689 (mile 0) and Y_{Rich} = 76,339 (mile 20.04) in MSA category 2 urban areas. Run the R scripts *Figure6A* and *Figure6B* to replicate these figures.

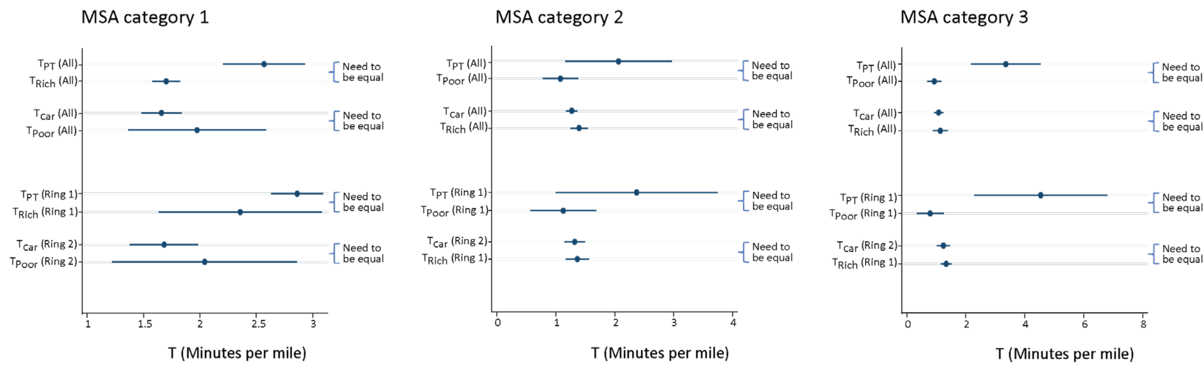


Figure 7. **Comparison of $\widehat{\boldsymbol{T}}_{\boldsymbol{PT},\boldsymbol{j}}$ and $\widehat{\boldsymbol{T}}_{\boldsymbol{Car},\boldsymbol{j}}$ (estimates of model 2.1) to $\widehat{\boldsymbol{T}}_{\boldsymbol{Poor},\boldsymbol{j}}$ and $\widehat{\boldsymbol{T}}_{\boldsymbol{Rich},\boldsymbol{j}}$ (estimates of model 2.2) across each MSA category of urban areas and their first two rings.** In Table 2 I indicate the transportation technology assignments that make MCCT model bid-rent curves consistent with observed HH income patterns. Here I determine if estimated $\widehat{T}_{PT,j}$, $\widehat{T}_{Car,j}$, $\widehat{T}_{Poor,j}$, and $\widehat{T}_{Rich,j}$ (dots with 5th and 95th confidence interval represented by the line through the dots) allow for the matches that align theory with observed data. Matches that need to be equal to align theory with observed data are indicated for each MSA category (All) and for each set of rings in each MSA (Rings 1 and 2). The ring analysis of MSA category 1 is the only case where $\widehat{T}_{PT,j}$, $\widehat{T}_{Car,j}$, $\widehat{T}_{Poor,j}$, and $\widehat{T}_{Rich,j}$ align theory with observed data. An 'All' comparison in MSA category 1 urban areas and all other 'All' and 'Ring' comparisons suggest that the representative poor and rich commuter are not using the commute technology that aligns theory with observed data.

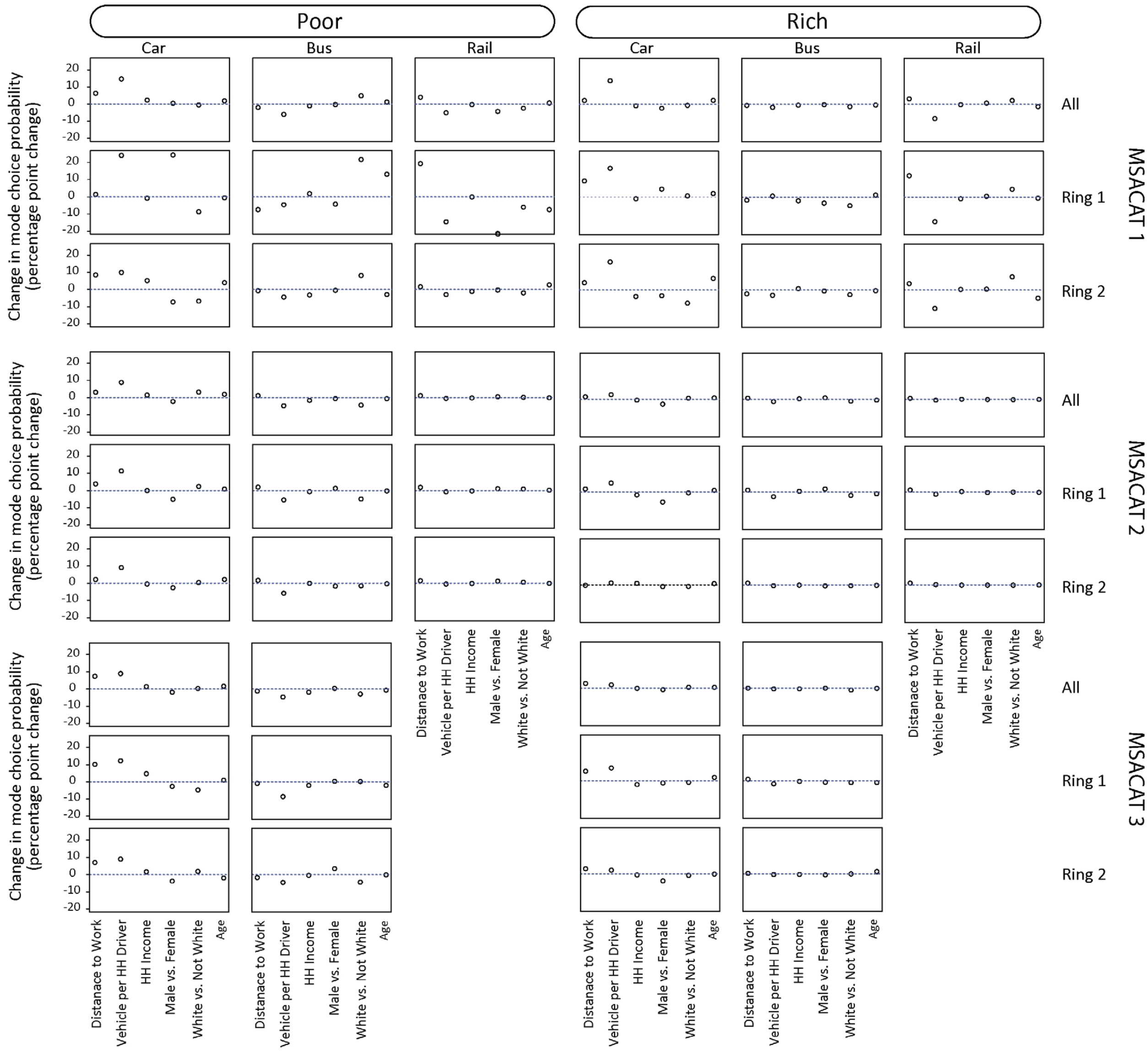


Figure 8. Simulated impacts of 1 standard deviation changes explanatory variables on mode choice in mode choice model (2.4). Each sub-graph gives a mean commuter's predicted percentage point change in car, bus, or rail probability use (y-axis) given a *ceteris paribus* 1 SD or status change in each of their independent variable values, indicated along the x-axis. All simulations are based on regression results that are weighted with 2017 NHTS person-level weights. I do not show the change in $\widehat{P}_{j,Inc,Geog,MSA}$ for *j* = 'other,' 'walk,', and 'bike.'

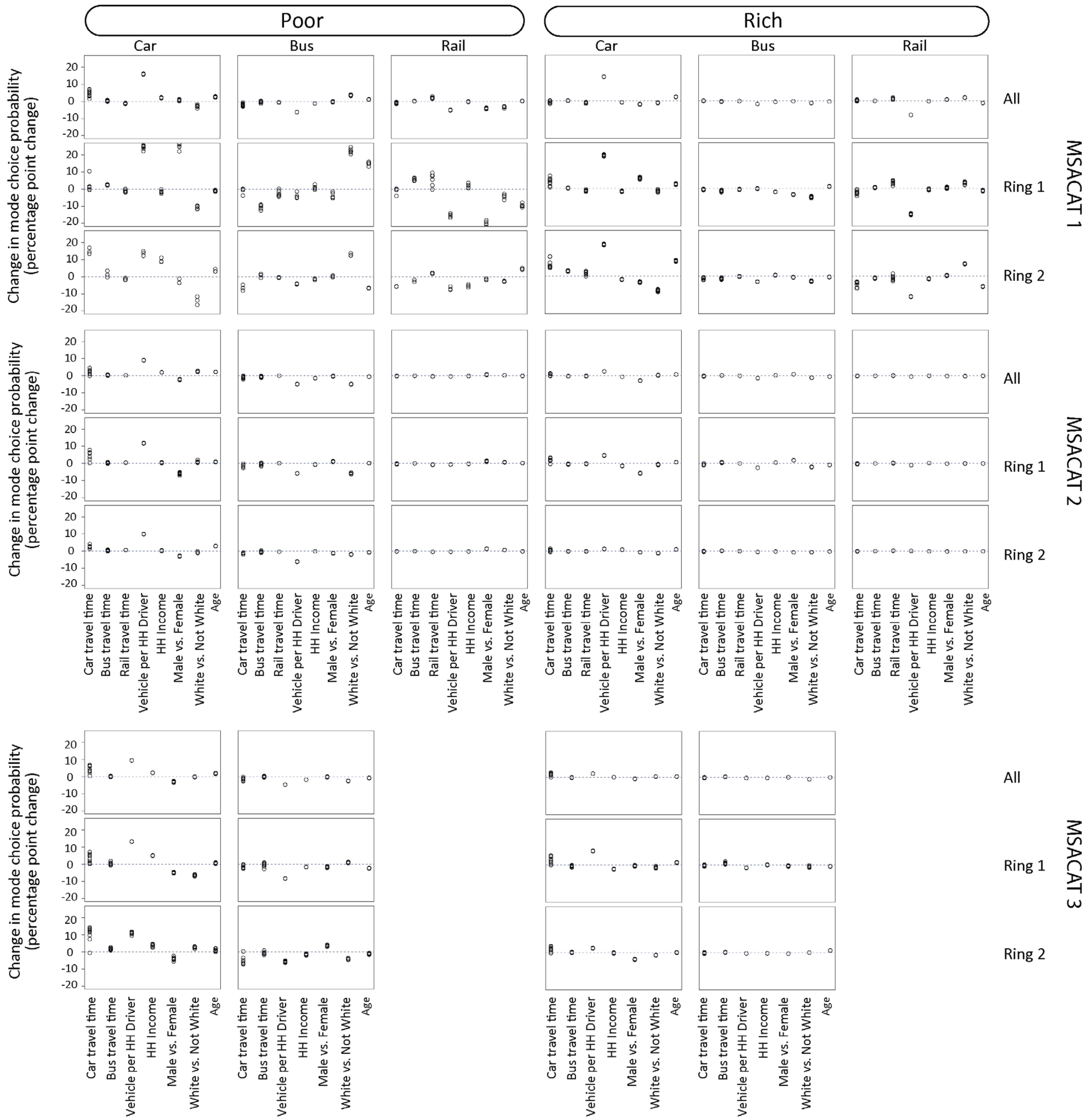


Figure 9. Simulated impacts of 1 standard deviation changes explanatory variables on mode choice in mode choice model (2.5). Each sub-graph gives a mean commuter's predicted percentage point change in car, bus, or rail probability use (y-axis) given a *ceteris paribus* 1 SD or status change in each of their independent variable values, indicated along the x-axis. There are 10 predicted percentage point changes for each *j*,*Inc*,*Geog*,*MSA* because there are 10 sets of commute times by mode (T_{ij}) for each commuter. In many cases 10 distinct points for each *j*,*Inc*,*Geog*,*MSA* cannot be see because the points are on top of each other. All simulations are based on regression results that are weighted with 2017 NHTS person-level weights. I do not show the change in $\hat{p}_{j,Inc,Geog,MSA}$ for *j* = 'other,' 'walk,', and 'bike.'

Tables

Table 1. Estimates of model (2.1)

MSA category	Pop. den. of home CT	Ring	T_{Car}	T_{PT}	F	Mean miles to work	Mean minutes to work
1	All		1.659 (0.082)	2.569 (0.169)	14.04 (1.484)	7.72	27.24
	> 25,000	1st	2.267 (0.495)	2.862 (0.103)	7.712 (2.908)	6.04	35.10
	10,000 – 24,999	2nd	1.683 (0.140)	3.112 (0.348)	7.081 (3.241)	7.34	30.07
2	All		1.270 (0.046)	2.064 (0.447)	14.60 (2.680)	8.20	22.14
	> 4,000	1st	1.297 (0.059)	2.373 (0.677)	13.87 (3.363)	7.63	23.31
	2,000 – 4,000	2nd	1.322 (0.087)	2.239 (0.385)	7.262 (4.730)	8.42	21.57
3	All		1.079 (0.067)	3.358 (0.509)	8.479 (2.921)	7.08	18.80
	> 4,000	1st	1.244 (0.170)	4.543 (0.978)	6.642 (4.155)	5.49	19.18
	2,000 – 4,000	2nd	1.239 (0.094)	5.296 (1.163)	-4.176 (3.918)	6.29	18.21

Notes: F is PT wait and egress time relative to car wait and egress time or PT wait and egress time less car wait and egress time. All regression results are weighted with 2017 NHTS person-level weights. Weights reflect the probability of a household or person being selected for survey participation and survey nonresponse. The weights can be used to produce population estimates at the national, census-region, or MSA category levels. See https://nhts.ornl.gov/assets/2017%20NHTS%20Weighting%20Report.pdf for more details. In this case results are given at the MSA category level: 1) MSA of 1 million or more, with heavy rail PT; 2) MSA of 1 million or more, but no heavy rail; 3) MSA less than 1 million; and 4) not in a MSA. I only include commuters that live within 20 miles of their work. MSA category 1 and 2 standard errors are clustered at CBSA level. MSA category 3 standard errors are clustered at CDIVMSAR level. ‘Car’ includes the technologies Car, SUV, Van, Pickup truck, Golf cart / Segway, and Motorcycle / Moped, and RV (motor home, ATV, snowmobile). ‘Bus’ includes the technologies School bus, Public or Commuter bus, Paratransit / Dial-a-ride, Private / Charter / Tour / Shuttle bus, and City-to-city bus (Greyhound, Megabus). ‘Rail’ includes Amtrak / Commuter rail and Subway / Elevated / Light rail / Street car. ‘Other’ includes Taxi / Limo (including Uber / Lyft), Rental car (Including Zipcar / Car2Go), Airplane, Boat / Ferry / Water taxi, and Something Else. If a commuter uses more than one mode they were asked to select their primary mode.

Table 2. Comparing observed urban areas to theoretical urban areas

MSA category	Observed ring pattern order	LHS of inequality (1.4) should be…	Transportation speeds (see Table 1)	LHS of inequality (1.4) is…
1	Rich, Poor	< 0.75	All, $T_{PT} = T_{Poor}$, $T_{Car} = T_{Rich}$	1.78
			All, $T_{Car} = T_{Poor}$, $T_{PT} = T_{Rich}$	-2.12*
			Rings, $T_{PT} = T_{Poor}$, $T_{Car} = T_{Rich}$	1.42
			Rings, $T_{Car} = T_{Poor}$, $T_{PT} = T_{Rich}$	-2.77*
2	Poor, Rich	> 0.75	All, $T_{PT} = T_{Poor}$, $T_{Car} = T_{Rich}$	0.96*
			Rings, $T_{PT} = T_{Poor}$, $T_{Car} = T_{Rich}$	1.06*
3	Poor, Rich	> 0.75	All, $T_{PT} = T_{Poor}$, $T_{Car} = T_{Rich}$	1.49*
			Rings, $T_{PT} = T_{Poor}$, $T_{Car} = T_{Rich}$	1.58*

Notes: Like Glaeser et al. (2008) I assume $\varepsilon_Y^A = [0.25, 0.50]$ and $\varepsilon_Y^W = 0.75$. Further, Y_{Poor} = 61,681 and Y_{Rich} = 78,968 in MSA 1 category urban areas and Y_{Poor} = 39,689 and Y_{Rich} = 76,339 in MSA category 2 and 3 urban areas. *MCCT model bid-rent curves are consistent with observed ring pattern of HH incomes in the given MSA category.

Table 3. Estimates of model (2.2)

MSA category	Pop. den. of home CT	Ring	T_{Rich}	T_{Poor}	F	Mean miles to work		Mean minutes to work	
						Poor	Rich	Poor	Rich
1	All		1.702 (0.057)	1.975 (0.282)	2.059 (0.963)	6.77	8.04	28.08	27.07
	> 25,000	1st	2.358 (0.324)	3.434 (0.325)	-3.285 (1.509)	5.85	5.93	36.08	33.78
	10,000 – 24,999	2nd	1.944 (0.080)	2.042 (0.379)	4.193 (2.580)	6.75	7.71	33.14	29.55
2	All		1.395 (0.073)	1.080 (0.147)	7.254 (2.200)	6.73	8.75	24.29	21.64
	> 4,000	1st	1.512 (0.116)	1.126 (0.277)	8.636 (3.624)	6.66	8.00	26.53	22.37
	2,000 – 4,000	2nd	1.365 (0.098)	1.260 (0.218)	4.581 (3.042)	6.24	9.27	21.14	21.59
3	All		1.138 (0.109)	0.932 (0.105)	3.903 (1.978)	6.10	7.82	20.05	18.98
	> 4,000	1st	1.602 (0.140)	0.791 (0.197)	6.884 (3.598)	4.97	6.50	21.14	20.88
	2,000 – 4,000	2nd	1.334 (0.077)	1.144 (0.206)	5.261 (1.809)	6.13	6.64	20.36	17.12

Notes: *F* in this case is the estimated wait and egress time of poor commuters less the wait and egress time of rich commuters. All regression results are weighted with 2017 NHTS person-level weights. Weights reflect the probability of a household or person being selected for survey participation and survey nonresponse. The weights can be used to produce population estimates at the national, census-region, or MSA category levels. See https://nhts.ornl.gov/assets/2017%20NHTS%20Weighting%20Report.pdf for more details. In this case results are given at the MSA category level: 1) MSA of 1 million or more, with heavy rail PT; 2) MSA of 1 million or more, but no heavy rail; 3) MSA less than 1 million; and 4) not in a MSA. I only include commuters that live within 20 miles of their work. MSA category 1 and 2 standard errors are clustered at CBSA level. MSA category 3 standard errors are clustered at CDIVMSAR level.

Table 4. Percentage breakdown of commute mode choices in 2017 NHTS survey. Commute mode choice is broken down by HH income category, MASA category, and population density in respondents' home CTs.

MSA category	**Pop. den. in respondents' home CTs**	**Ring**	**Poor HH**				**Middle HH**				**Rich HH**			
			Car	**Bus**	**Rail**	**Total commuters**	**Car**	**Bus**	**Rail**	**Total commuters**	**Car**	**Bus**	**Rail**	**Total commuters**
1	All		68.3%	12.5%	8.3%	3,225,116	75.9%	7.6%	9.0%	9,765,859	73.9%	3.6%	14.4%	15,832,495
	> 25,000	1st	23.0%	30.8%	32.5%	572,592	27.8%	18.6%	35.5%	1,705,277	22.5%	7.4%	45.6%	2,781,433
	10,000 – 24,999	2nd	66.6%	16.3%	6.4%	976,194	73.6%	10.3%	6.8%	2,335,146	63.5%	7.0%	21.0%	3,008,043
2	All		80.3%	10.2%	0.9%	4,212,902	91.0%	3.7%	0.8%	12,370,690	92.4%	2.4%	1.2%	13,253,090
	> 4,000	1st	73.6%	11.9%	1.5%	2,294,803	85.6%	6.7%	1.3%	5,755,535	86.8%	4.5%	2.1%	5,543,108
	2,000 – 4,000	2nd	86.6%	11.3%	0.2%	1,011,919	94.0%	1.3%	0.6%	3,239,888	95.1%	1.1%	1.3%	3,707,939
3	All		83.1%	7.3%	0.0%	5,582,938	93.6%	2.1%	0.0%	15,104,700	95.9%	0.9%	0.0%	11,167,180
	> 4,000	1st	73.0%	13.9%	0.0%	1,813,960	87.4%	3.8%	0.0%	3,299,378	89.2%	3.0%	0.0%	1,781,587
	2,000 – 4,000	2nd	87.0%	6.5%	0.0%	1,474,519	93.6%	1.9%	0.0%	3,376,685	95.1%	0.6%	0.0%	2,454,012

Notes: Mode choice distribution uses 2017 NHTS person-level weights. Weights reflect the probability of a person being selected for survey participation and survey nonresponse. The weights can be used to produce population estimates at the national, census-region, or MSA category levels. In this case results are given at the MSA category level: 1) MSA of 1 million or more, with heavy rail; 2) MSA of 1 million or more, but no heavy rail; 3) MSA less than 1 million; and 4) not in MSA.

Table 5. Mean and standard deviation of independent variables in mode choice models (2.4) and (2.5) for poor and rich commuters by location and geographic unit.

Geographic unit	**All**		**1st Ring**		**2nd Ring**	
	Poor	**Rich**	**Poor**	**Rich**	**Poor**	**Rich**
	MSACAT = 1					
Time to work given chosen mode (minutes)	27.11 (25.92)	25.86 (21.65)	39.12 (41.72)	32.33 (21.73)	26.59 (23.48)	26.91 (21.58)
Distance (miles)	7.11 (5.40)	8.28 (5.55)	6.53 (5.17)	5.69 (4.83)	6.88 (5.37)	7.83 (5.36)
Vehicle per Driver	0.85 (0.48)	1.07 (0.5)	0.48 (0.51)	0.57 (0.49)	0.89 (0.46)	0.99 (0.47)
HH Income (USD)	18055 (9964)	153618 (62662)	15647 (9316)	155409 (66354)	18850 (10217)	150243 (63267)
Male (percentage)	45.95	49.68	41.07	50.56	48.59	53.82
White (percentage)	50.18	77.34	32.14	77.69	47.42	70.49
Age	40.81 (14.86)	45.48 (13.94)	40.12 (15.1)	41.84 (13.29)	40.50 (14.51)	43.46 (13.17)
	MSACAT = 2					
Time to work given chosen mode (minutes)	23.59 (27.29)	21.81 (18.13)	25.16 (25.91)	22.33 (18.33)	23.95 (23.58)	21.40 (17.04)
Distance (miles)	7.20 (5.30)	8.96 (5.43)	7.09 (5.2)	8.53 (5.34)	7.15 (5.22)	9.03 (5.39)
Vehicle per Driver	0.92 (0.43)	1.19 (0.47)	0.89 (0.43)	1.15 (0.44)	0.92 (0.43)	1.18 (0.5)
HH Income (USD)	17121 (9427)	141124 (58583)	16937 (9361)	136637 (58238)	17379 (9404)	145048 (59263)
Male (percentage)	46.17	51.1	47.63	50.94	45.76	52.81
White (percentage)	58.59	84.24	55.35	82.26	56.26	84.44
Age	39.90 (15.17)	45.89 (13.77)	38.67 (14.66)	44.89 (13.58)	39.82 (14.98)	46.6 (13.87)
	MSACAT = 3					
Time to work given chosen mode (minutes)	19.39 (26.19)	17.88 (17.35)	21.19 (35.85)	16.80 (20.36)	19.06 (30.13)	16.72 (15.81)
Distance (miles)	6.05 (4.96)	7.65 (5.15)	4.78 (4.35)	5.65 (4.58)	5.16 (4.21)	6.50 (4.62)
Vehicle per Driver	0.95 (0.51)	1.27 (0.57)	0.84 (0.48)	1.19 (0.52)	0.93 (0.50)	1.22 (0.52)
HH Income (USD)	17060 (9330)	130182 (53078)	15818 (9207)	124301 (51377)	16684 (9684)	128726 (54341)
Male (percentage)	44.3	50.32	44.04	49.17	44.18	50.75
White (percentage)	67.48	89.34	65.65	84.69	60.14	88.35
Age	40.11 (15.67)	47.49 (13.87)	37.19 (14.44)	46.69 (14.09)	40.38 (15.76)	46.94 (14.02)

Notes: The data above is not weighted using 2017 NHTS person-level weights. Data can be recreated by running R scripts *MSACAT1PoorwithDistStandardize* and *MSACAT1RichwithDistStandardize* with the dataset *MSACAT1.csv*; *MSACAT2PoorwithDistStandardize* and *MSACAT2RichwithDistStandardize* with the dataset *MSACAT2.csv*; and *MSACAT3PoorwithDistStandardize* and *MSACAT3RichwithDistStandardize* with the dataset *MSACAT3.csv*. 'Time to work' means and standard deviations are a function of the modes that each commuter chose in 2017.

Table 6. Estimates of mode choice model (2.4) at the mean commuter's set of explanatory variables

			Poor HH					Rich HH				
MSA category	**Pop. den. in respondents' home CTs**	**Ring**	**Car**	**Bus**	**Rail**	**Bike**	**Walk**	**Car**	**Bus**	**Rail**	**Bike**	**Walk**
1	All		71.3	11.6	7.5	0.9	7.3	79.3	2.9	11.3	1.6	3.6
	> 25,000	1st	23.4	26.8	35.2	0.0	13.8	23.7	7.2	43.6	5.6	17.0
	10,000 – 25,000	2nd	74.6	10.2	4.1	2.0	7.1	67.4	6.6	18.2	2.6	3.4
2	All		85.5	7.0	0.8	1.1	5.2	93.4	2.0	1.0	1.3	1.5
	> 4,000	1st	80.8	8.1	1.4	1.9	6.9	88.3	3.7	2.0	2.5	2.5
	2,000 – 4,000	2nd	85.5	7.7	0.8	0.8	4.5	96.6	0.9	0.5	0.9	0.5
3	All		84.8	6.6		1.9	5.4	96.0	0.7		1.2	1.5
	> 4,000	1st	72.5	12.0		2.6	10.5	89.1	2.1		3.5	4.8
	2,000 – 4,000	2nd	85.4	6.9		2.6	4.3	95.4	0.8		1.7	1.8

Notes. A mean commuter is defined for each MSA category, income group, and geographic area (all urban area, ring 1, or ring 2) combination. Predicted probabilities for the commute mode categories of bike, other, and walk are not reported in this table. Data can be recreated by running R scripts *MSACAT1PoorwithDistStandardize* and *MSACAT1RichwithDistStandardize* with the dataset *MSACAT1.csv*; *MSACAT2PoorwithDistStandardize* and *MSACAT2RichwithDistStandardize* with the dataset *MSACAT2.csv*; and *MSACAT3PoorwithDistStandardize* and *MSACAT3RichwithDistStandardize* with the dataset *MSACAT3.csv*.

Table 7. Estimates of mode choice mode (2.5) at the mean commuter's set of explanatory variables

			Poor HH					Rich HH				
MSA category	**Pop. den. in respondents' home CTs**	**Ring**	**Car**	**Bus**	**Rail**	**Bike**	**Walk**	**Car**	**Bus**	**Rail**	**Bike**	**Walk**
1	All		70.7 (70.5, 71.0)	11.9 (11.8, 12.0)	7.4 (7.2, 7.6)	0.9 (0.8, 1.0)	7.6 (7.3, 7.7)	79.3 (79.2, 79.4)	2.9 (2.9, 2.9)	11.3 (11.3, 11.4)	1.6 (1.6, 1.7)	3.5 (3.4, 3.5)
	> 25,000	1st	23.6 (23.4, 24.0)	26.7 (25.9, 27.0)	34.3 (33.7, 34.7)	0.0 (0.0, 0.0)	14.5 (14.0, 15.0)	24.1 (23.9, 24.2)	7.1 (7.1, 7.3)	44.3 (43.9, 44.5)	5.5 (5.4, 5.7)	16.2 (15.8, 16.4)
	10,000 – 25,000	2nd	67.3 (65.6, 68.3)	13.8 (13.5, 14.0)	10.9 (9.9, 11.6)	0.6 (0.5, 0.7)	6.2 (5.2, 7.4)	65.6 (65.1, 65.9)	6.6 (6.5, 6.7)	19.7 (19.5, 20)	2.7 (2.5, 3.0)	3.5 (2.9, 3.9)
2	All		85.5 (85.4, 85.6)	7.1 (7.0, 7.2)	0.7 (0.7, 0.8)	1.1 (1.0, 1.1)	5.1 (5.0, 5.1)	93.3 (93.3, 93.4)	2.0 (2.0, 2.0)	0.9 (0.9, 1.0)	1.3 (1.3, 1.4)	1.6 (1.6, 1.6)
	> 4,000	1st	80.2 (80.1, 80.4)	8.4 (8.3, 8.6)	1.2 (1.1, 1.3)	1.9 (1.8, 1.9)	7.2 (7.1, 7.4)	88.1 (88, 88.2)	3.6 (3.6, 3.7)	1.9 (1.9, 1.9)	2.6 (2.5, 2.6)	2.7 (2.7, 2.8)
	2,000 – 4,000	2nd	84.8 (84.7, 85.0)	8.0 (7.9, 8.1)	0.8 (0.7, 0.8)	0.8 (0.7, 0.8)	4.8 (4.7, 5.0)	96.3 (96.2, 96.4)	0.9 (0.9, 1.0)	0.6 (0.6, 0.6)	0.8 (0.8, 0.9)	0.8 (0.7, 0.9)
3	All		85.2 (85.0, 85.3)	6.5 (6.4, 6.6)		1.9 (1.8, 1.9)	5.1 (5.1, 5.2)	96.2 (96.1, 96.2)	0.7 (0.7, 0.8)		1.1 (1.1, 1.2)	1.4 (1.4, 1.4)
	> 4,000	1st	74.5 (74.3, 74.7)	11.6 (11.5, 11.7)		2.6 (2.6, 2.7)	8.6 (8.5, 8.8)	89.6 (89.2, 89.9)	2.1 (2.0, 2.2)		3.5 (3.4, 3.7)	4.4 (4.3, 4.6)
	2,000 – 4,000	2nd	84.0 (83.0, 86.5)	8.1 (7.0, 8.7)		3.4 (2.5, 4.2)	3.4 (3.1, 3.7)	95.3 (95, 95.6)	0.8 (0.8, 0.9)		1.6 (1.4, 1.9)	1.9 (1.8, 1.9)

Table 8. Los Angeles-2003

	Control Group	***Historic transit network***					
	Dep. variable	**Median HH income**		**Per capita income**		**Poverty rate**	
	Spatial unit	**CT**	**BG**	**CT**	**BG**	**CT**	**BG**
	$\widehat{\boldsymbol{\delta}}_{\mathbf{1980}}$	-4,911		-5,950***		0.041***	
		(3,202)		(2,264)		(0.012)	
Pre-treatment	$\widehat{\boldsymbol{\delta}}_{\mathbf{1990}}$	-2,065	-3,596	-3,104	-3,514	0.024**	0.047***
		(3,202)	(3,643)	(2,264)	(2,257)	(0.012)	(0.015)
	$\widehat{\boldsymbol{\delta}}_{\mathbf{2000}}$	0	0	0	0	0	0
	$\widehat{\boldsymbol{\delta}}_{\mathbf{2010}}$	3,922	7,916**	1,749	5,033**	0.019	0.029*
Post-treatment		(3,202)	(3,643)	(2,264)	(2,257)	(0.012)	(0.015)
	$\widehat{\boldsymbol{\delta}}_{\mathbf{2017}}$	7,027**	10,266***	1,029	3,553	0.014	0.031**
		(3,202)	(3,643)	(2,264)	(2,257)	(0.012)	(0.015)
	N	355	284	355	284	355	284
	$\mathbf{R^2}$	0.941	0.948	0.939	0.952	0.921	0.889
	Control Group	***Treated area buffers***					
	$\widehat{\boldsymbol{\delta}}_{\mathbf{1980}}$	1,714		2,674 (2,971)		0.003	
		(3,690)				(0.013)	
Pre-treatment	$\widehat{\boldsymbol{\delta}}_{\mathbf{1990}}$	-1,259	-2,861	252 (2,929)	-150	0.005	0.037
		(3,638)	(4,536)		(3,138)	(0.013)	(0.024)
	$\widehat{\boldsymbol{\delta}}_{\mathbf{2000}}$	0	0	0	0	0	0
	$\widehat{\boldsymbol{\delta}}_{\mathbf{2010}}$	-5,043	1,489	-62.00	4,600	0.021*	0.027
Post-treatment		(3,638)	(4,536)	(2,929)	(3,138)	(0.013)	(0.024)
	$\widehat{\boldsymbol{\delta}}_{\mathbf{2017}}$	1,217	5,158	539	4,035	-0.005	0.008
		(3,638)	(4,536)	(2,929)	(3,138)	(0.013)	(0.024)
	N	119	96	119	96	119	96
	$\mathbf{R^2}$	0.945	0.930	0.947	0.944	0.961	0.867

Notes. Treated areas are centered on Gold light rail line stations that opened in 2003. Historic transit network control group contains 59 half-mile radius circles around 59 randomly selected points on the rapid transit lines that existed or were planned as of 1925. The paper map of the existing and planned lines is from Kelker, De Leuw & Co. (1925). Nominal dollar amounts are inflated to 2017 USD using the “All items in Los Angeles-Long Beach-Anaheim, CA, all urban consumers, not seasonally adjusted” CPI (series ID: CUURS49ASA0, CUUSS49ASA0).

Table 9. Los Angeles-2005

	Control group	*Historic transit network*					
	Dep. variable	Median HH income		Per capita income		Poverty rate	
	Spatial unit	CT	BG	CT	BG	CT	BG
Pre-treatment	$\hat{\delta}_{1980}$	6,863**		716.0		-0.011	
		(3,088)		(1,958)		(0.014)	
	$\hat{\delta}_{1990}$	3,345	2,645	607	561	-0.016	-0.024
		(3,088)	(3,398)	(1,958)	(1,918)	(0.014)	(0.016)
	$\hat{\delta}_{2000}$	0	0	0	0	0	0
Post-treatment	$\hat{\delta}_{2010}$	11,105***	11,714***	1,536 (1,958)	2,281	0.046***	0.042***
		(3,088)	(3,398)		(1,918)	(0.014)	(0.016)
	$\hat{\delta}_{2017}$	7,023**	8,717**	121	-1,266	0.014	0.007
		(3,088)	(3,398)	(1,958)	(1,918)	(0.014)	(0.016)
	N	365	292	365	292	365	292
	R^2	0.930	0.947	0.937	0.957	0.857	0.843
	Control group	*Treated area buffers*					
Pre-treatment	$\hat{\delta}_{1980}$	68.84		-1,562		-0.005	
		(3,529)		(1,440)		(0.024)	
	$\hat{\delta}_{1990}$	-310	-1,796	-1,546	-1,604	-0.001	-0.008
		(3,529)	(3,891)	(1,440)	(1,708)	(0.024)	(0.026)
	$\hat{\delta}_{2000}$	0	0	0	0	0	0
Post-treatment	$\hat{\delta}_{2010}$	1,051	4,474	-769	1,524	0.010	0.003
		(3,529)	(3,891)	(1,440)	(1,708)	(0.024)	(0.026)
	$\hat{\delta}_{2017}$	-590	399	-11.0	-41.0	0.011	0.0003
		(3,529)	(3,891)	(1,440)	(1,708)	(0.024)	(0.026)
	N	140	112	140	112	140	112
	R^2	0.816	0.851	0.838	0.824	0.569	0.563

Notes. Treated areas are centered on orange rapid bus line stations that opened in 2005-2006. Historic transit network control group contains 59 half-mile radius circles around 59 randomly selected points on the rapid transit lines that existed or were planned as of 1925. The paper map of the existing and planned lines is from Kelker, De Leuw & Co. (1925). Nominal dollar amounts are inflated to 2017 USD using the "All items in Los Angeles-Long Beach-Anaheim, CA, all urban consumers, not seasonally adjusted" CPI (series ID: CUURS49ASA0, CUUSS49ASA0).

Table 10. Minneapolis-2004

	Control group	*Historic transit network*					
	Dep. variable	**Median HH income**		**Per capita income**		**Poverty rate**	
	Spatial unit	**CT**	**BG**	**CT**	**BG**	**CT**	**BG**
	$\hat{\delta}_{1980}$	12,556***		1,375		0.024	
		(4,389)		(2,237)		(0.023)	
Pre-treatment	$\hat{\delta}_{1990}$	-3,823	-1,608	2,226	1,461	0.065***	0.109***
		(4,389)	(3,522)	(2,237)	(2,328)	(0.023)	(0.029)
	$\hat{\delta}_{2000}$	0	0	0	0	0	0
	$\hat{\delta}_{2010}$	-8,382*	-1,553	2,525	3,656	0.049**	0.107***
Post-treatment		(4,389)	(3,522)	(2,237)	(2,328)	(0.023)	(0.029)
	$\hat{\delta}_{2017}$	4,888	9,913***	-2,134	-1,619	-0.019	0.024
		(4,527)	(3,522)	(2,237)	(2,328)	(0.023)	(0.029)
	N	363	290	363	292	363	292
	R^2	0.820	0.922	0.852	0.869	0.860	0.794
	Control group	***Areas that will be treated in the future***					
	$\hat{\delta}_{1980}$	15,374*		10,398**		0.029	
		(7,812)		(4,441)		(0.038)	
Pre-treatment	$\hat{\delta}_{1990}$	-7,493	3,789	6,655	6,627	0.077**	0.010**
		(7,812)	(6,370)	(4,441)	(4,867)	(0.038)	(0.049)
	$\hat{\delta}_{2000}$	0	0	0	0	0	0
	$\hat{\delta}_{2010}$	-7,620	-150.5	-1,833	-1,045	0.101***	0.154***
Post-treatment		(7,812)	(6,370)	(4,441)	(4,867)	(0.038)	(0.049)
	$\hat{\delta}_{2017}$	11,157	12,632*	-3,101	-4,263	0.008	0.036
		(7,969)	(6,370)	(4,441)	(4,867)	(0.038)	(0.049)
	N	163	130	163	132	163	132
	R^2	0.655	0.869	0.828	0.839	0.796	0.692
	Control group	***Treated area buffers***					
	$\hat{\delta}_{1980}$	-3,453		546.3		-0.046	
		(9,061)		(3,691)		(0.042)	
Pre-treatment	$\hat{\delta}_{1990}$	563.8	4,593	3,291	3,682	-0.026	-0.002
		(9,061)	(5,393)	(3,691)	(4,409)	(0.042)	(0.055)
	$\hat{\delta}_{2000}$	0	0	0	0	0	0
	$\hat{\delta}_{2010}$	-1,712	3,437	3,145	353.5	-0.034	0.037
Post-treatment		(9,061)	(5,393)	(3,691)	(4,409)	(0.042)	(0.055)
	$\hat{\delta}_{2017}$	-1,940	10,898**	86.0	2,028	-0.037	0.0243
		(9,284)	(5,393)	(3,691)	(4,409)	(0.042)	(0.055)
	N	165	134	168	136	167	136
	R^2	0.386	0.785	0.696	0.738	0.743	0.630

Notes. Treated areas are centered on Blue light rail line stations that opened in 2004. The historic network control group includes 56 half-mile radius circles centered on randomly selected points on its historic street car track network that existed as of the early 1950s. Nominal dollar amounts are inflated to 2017 USD using the "All items in Minneapolis-St. Paul-Bloomington, MN-WI, all urban consumers, not seasonally adjusted" CPI (series ID: CUURS24ASA0, CUUSS24ASA0).

Table 11. Denver-1994

	Control group	*Historic transit network*					
	Dep. variable	**Med. HH income**		**Per capita income**		**Poverty rate**	
	Spatial unit	**CT**	**BG**	**CT**	**BG**	**CT**	**BG**
	$\hat{\delta}_{1980}$	-1,493		-2,536		0.052**	
Pre-		(4,406)		(3,682)		(0.024)	
treatment	$\hat{\delta}_{1990}$	-2,373	-4,080	-2,534		0.062***	0.092***
		(4,406)	(4,566)	(3,682)	-6,325 (3,837)	(0.024)	(0.030)
	$\hat{\delta}_{2000}$	0	0	0	0	0	0
Post-	$\hat{\delta}_{2010}$	7,040		347	-2,913	-0.039*	-0.031
treatment		(4,406)	2,237 (4,566)	(3,682)	(3,837)	(0.024)	(0.030)
	$\hat{\delta}_{2017}$	3,714	-572	696	-1,204	-0.063***	-0.054*
		(4,406)	(4,566)	(3,682)	(3,837)	(0.024)	(0.030)
	N	145	116	145	116	145	116
	R^2	0.925	0.948	0.915	0.933	0.908	0.884
	Control group	***Areas that will be treated in the future***					
	$\hat{\delta}_{1980}$	-5,927		-6,932**		0.071***	
Pre-		(5,191)		(3,210)		(0.023)	
treatment	$\hat{\delta}_{1990}$	-8,345	-5,088	-6,802**	-4,622	0.096***	0.106***
		(5,191)	(5,512)	(3,210)	(5,665)	(0.023)	(0.036)
	$\hat{\delta}_{2000}$	0	0	0	0	0	0
Post-	$\hat{\delta}_{2010}$	18,561***	18,307***	8,622***	13,433**	-0.073***	-0.098***
treatment		(5,191)	(5,512)	(3,210)	(5,665)	(0.023)	(0.036)
	$\hat{\delta}_{2017}$	29,427***	23,136***	12,660***	17,853***	-0.110***	-0.118***
		(5,191)	(5,512)	(3,210)	(5,665)	(0.023)	(0.036)
	N	130	104	130	104	130	104
	R^2	0.923	0.929	0.888	0.809		
	Control group	***Treated area buffers***					
	$\hat{\delta}_{1980}$	-405.8		-3,169		0.012	
Pre-		(3,809)		(3,239)		(0.027)	
treatment	$\hat{\delta}_{1990}$	1,523	6,579	-733.7	-1,384	0.002	0.0004
		(3,809)	(4,829)	(3,239)	(4,044)	(0.027)	(0.039)
	$\hat{\delta}_{2000}$	0	0	0	0	0	0
Post-	$\hat{\delta}_{2010}$	2,593	3,628	-4,561	-5,310	-0.011	0.004
treatment		(3,809)	(4,829)	(3,239)	(4,044)	(0.027)	(0.039)
	$\hat{\delta}_{2017}$	1,714	3,285	-2,360	-1,930	0.0047	-0.001
		(3,809)	(4,829)	(3,239)	(4,044)	(0.027)	(0.039)
	N	100	80	100	80	100	80
	R^2	0.898	0.837	0.910	0.861	0.893	0.846

Notes. Treated areas are centered on D light rail line stations that opened in 1994. The historic transit network control group includes 19 half-mile radius circles centered on randomly selected points on its historic street car track network that existed as of 1936. Nominal dollar amounts are inflated to 2017 USD using the "All items in Denver-Aurora-Lakewood, CO, all urban consumers, not seasonally adjusted" CPI (series ID: CUURS49ASA0, CUUSS49ASA0).

Table 12. Denver-2006

	Control group	*Historic transit network*					
	Dep. Variable	**Med. HH income**		**Per capita income**		**Poverty rate**	
	Spatial unit	**CT**	**BG**	**CT**	**BG**	**CT**	**BG**
	$\hat{\boldsymbol{\delta}}_{\mathbf{1980}}$	9,217		-3,454		-0.016	
		(5,825)		(3,739)		(0.018)	
Pre-	$\hat{\boldsymbol{\delta}}_{\mathbf{1990}}$	10,187*	7,773			-0.062***	-0.058***
treatment		(5,825)	(6,327)	2,104 (3,739)	-2,299 (3,899)	(0.018)	(0.020)
	$\hat{\boldsymbol{\delta}}_{\mathbf{2000}}$	0	0	0	0	0	0
	$\hat{\boldsymbol{\delta}}_{\mathbf{2010}}$	-20,057***	-34,718***	-8,331**	-14,403***	0.020	0.032
Post-		(5,825)	(6,327)	(3,739)	(3,899)	(0.018)	(0.020)
treatment	$\hat{\boldsymbol{\delta}}_{\mathbf{2017}}$	-26,143***	-31,425***	-10,823***	-13,702***	0.046**	0.055***
		(5,825)	(6,327)	(3,739)	(3,899)	(0.018)	(0.020)
	N	160	128	160	128	160	128
	$\mathbf{R^2}$	0.890	0.899	0.885	0.909	0.875	0.849
	Control group	***Areas that will be treated in the future***					
	$\hat{\boldsymbol{\delta}}_{\mathbf{1980}}$	4,783		-7,850**		0.002	
		(6,587)		(3,491)		(0.016)	
Pre-	$\hat{\boldsymbol{\delta}}_{\mathbf{1990}}$	4,215	6,764	-2,165	-595.9	-0.027*	-0.044*
treatment		(6,587)	(7,210)	(3,491)	(5,404)	(0.016)	(0.025)
	$\hat{\boldsymbol{\delta}}_{\mathbf{2000}}$	0	0	0	0	0	0
	$\hat{\boldsymbol{\delta}}_{\mathbf{2010}}$	-8,536	-18,647**	-56.0	1,943	-0.013	-0.035
Post-		(6,587)	(7,210)	(3,491)	(5,404)	(0.016)	(0.025)
treatment	$\hat{\boldsymbol{\delta}}_{\mathbf{2017}}$	-430	-7,718	1,141	5,354	-0.002	-0.010
		(6,587)	(7,210)	(3,491)	(5,404)	(0.016)	(0.025)
	N	145	116	145	116	145	116
	$\mathbf{R^2}$	0.859	0.870	0.885	0.830	0.879	0.831
	Control group	***Treated area buffers***					
	$\hat{\boldsymbol{\delta}}_{\mathbf{1980}}$	1,025		-806		0.0007	
		(8,023)		(3,972)		(0.010)	
Pre-	$\hat{\boldsymbol{\delta}}_{\mathbf{1990}}$	910	-2,846	-541	-1,837	-0.0002	0.0032
treatment		(8,023)	(8,552)	(3,972)	(4,666)	(0.010)	(0.012)
	$\hat{\boldsymbol{\delta}}_{\mathbf{2000}}$	0	0	0	0	0	0
	$\hat{\boldsymbol{\delta}}_{\mathbf{2010}}$	-2,913	-15,346*	-713	-6,799	-0.0022	0.0035
Post-		(8,023)	(8,552)	(3,972)	(4,666)	(0.010)	(0.012)
treatment	$\hat{\boldsymbol{\delta}}_{\mathbf{2017}}$	19.5	-7,481	397	-1,984	0.0036	0.0068
		(8,023)	(8,552)	(3,972)	(4,666)	(0.010)	(0.012)
	N	130	104	130	104	130	104
	$\mathbf{R^2}$	0.754	0.781	0.810	0.763	0.827	0.822

Notes. Treated areas are half-mile radii circles around F line stations that opened in 2006. The historic transit network control group includes 19 half-mile radius circles centered on randomly selected points on its historic street car track network that existed as of 1936. Nominal dollar amounts are inflated to 2017 USD using the “All items in Denver-Aurora-Lakewood, CO, all urban consumers, not seasonally adjusted” CPI (series ID: CUURS49ASA0, CUUSS49ASA0).

Table 13. Phoenix-2008

	Control group	*Areas that will be treated in the future*					
	Dep. variable	**Med. HH income**		**Per capita income**		**Poverty rate**	
	Spatial unit	**CT**	**BG**	**CT**	**BG**	**CT**	**BG**
Pre-treatment	$\hat{\delta}_{1980}$	-4,718*		-1,281		0.043	
		(2,463)		(1,466)		(0.027)	
	$\hat{\delta}_{1990}$	-2,909	-31.18	-249.2	-4,757*	0.017	0.026
		(2,463)	(5,489)	(1,466)	(2,430)	(0.027)	(0.032)
	$\hat{\delta}_{2000}$	0	0	0	0		
Post-treatment	$\hat{\delta}_{2010}$	-276.9	-2,524	2,095	1,030 (2,430)	-0.005	0.015
		(2,463)	(5,489)	(1,466)		(0.027)	(0.032)
	$\hat{\delta}_{2017}$	7,228***	3,224	6,786***	5,595**	-0.052*	-0.016
		(2,463)	(5,489)	(1,466)	(2,430)	(0.027)	(0.032)
	N	285	228	285	228	285	228
	R^2	0.773	0.492	0.861	0.676	0.813	0.754
	Control group	*Treated area buffers*					
Pre-treatment	$\hat{\delta}_{1980}$	-980.2		-97.76		0.026	
		(2,266)		(1,453)		(0.025)	
	$\hat{\delta}_{1990}$	-1,837	4,426	177.0	-1,841	0.016	0.019
		(2,266)	(4,557)	(1,453)	(2,828)	(0.025)	(0.028)
	$\hat{\delta}_{2000}$	0	0	0	0	0	0
Post-treatment	$\hat{\delta}_{2010}$	1,056	1,040	358.7	3.617	-0.007	0.003
		(2,266)	(4,557)	(1,453)	(2,828)	(0.025)	(0.028)
	$\hat{\delta}_{2017}$	4,281*	3,211	3,831***	3,218	0.006	0.007
		(2,266)	(4,557)	(1,453)	(2,828)	(0.025)	(0.028)
	N	270	216	270	216	270	216
	R^2	0.836	0.620	0.881	0.640	0.816	0.814

Notes. Treated areas are half-mile radii circles around light rail stations that opened in 2008. Nominal dollar amounts are inflated to 2017 USD using the “All items in West - Size Class A, all urban consumers, not seasonally adjusted” CPI (series ID: CUURS400SA0, CUUSS400SA0).

Supplementary Information

SI Text 1.

In a city where the rich sort into the inner ring in equilibrium the following must hold,

$$\underbrace{-\frac{2W_{Rich}T}{A_{Rich}}}_{\substack{\text{Slope of rich}\\\text{bid-rent curve}}} < \underbrace{-\frac{2W_{Poor}T}{A_{Poor}}}_{\substack{\text{Slope of poor}\\\text{bid-rent curve}}} \tag{1.3}$$

$$\frac{W_{Rich}-W_{Poor}}{W_{Poor}} > \frac{A_{Rich}-A_{Poor}}{A_{Poor}} \tag{1.4}$$

$$\frac{W_{Rich}-W_{Poor}}{W_{Poor}}\frac{\overbrace{Y_{Poor}}^{\substack{\text{Poor HH}\\\text{income}}}}{\underbrace{Y_{Rich}}_{\substack{\text{Rich HH}\\\text{income}}}-Y_{Poor}} > \frac{A_{Rich}-A_{Poor}}{A_{Poor}}\frac{Y_{Poor}}{Y_{Rich}-Y_{Poor}} \tag{1.5}$$

$$\left(\frac{W_{Rich}-W_{Poor}}{Y_{Rich}-Y_{Poor}}\right)\frac{Y_{Poor}}{W_{Poor}} > \left(\frac{A_{Rich}-A_{Poor}}{Y_{Rich}-Y_{Poor}}\right)\frac{Y_{Poor}}{A_{Poor}} \tag{1.6}$$

$$\underbrace{\frac{\Delta W}{\Delta I}\times\frac{Y_{Poor}}{W_{Poor}}}_{\varepsilon_Y^W} > \underbrace{\frac{\Delta A}{\Delta I}\times\frac{Y_{Poor}}{A_{Poor}}}_{\varepsilon_Y^A} \tag{1.7}$$

In a city where the poor sort into the inner ring in equilibrium the following must hold,

$$-\frac{2W_{Rich}T_{Rich}}{A_{Rich}} > \underbrace{-\frac{2W_{Poor}T_{Poor}}{A_{Poor}}}_{\substack{\text{Poor HH bid-}\\\text{rent curve is}\\\text{steeper}}} \tag{1.8}$$

$$\frac{W_{Rich}T_{Rich}}{A_{Rich}} < \frac{W_{Poor}T_{Poor}}{A_{Poor}} \tag{1.9}$$

$$\frac{W_{Rich}T_{Rich}}{W_{Poor}T_{Poor}} < \frac{A_{Rich}}{A_{Poor}} \tag{1.10}$$

In the text I claim that if,

$$\frac{W_{Rich}T_{Rich}}{W_{Poor}T_{Poor}} < \frac{A_{Rich}}{A_{Poor}}$$

then

$$\varepsilon_Y^A + \frac{T_{Poor}-T_{Rich}}{T_{Poor}}\left(\frac{Y_{Poor}}{Y_{Rich}-Y_{Poor}} + \varepsilon_Y^W\right) > \varepsilon_Y^W$$

assuming the gap between the typical poor and rich area in an urban area, represented by $Y_{Rich} - Y_{Poor}$, is not too large. The following R script uses computation simulation to show that this if…then statement holds.

SI Text 2

Using data from the early 2000s, Glaeser et al. (2008) find that the typical poor commuter in a typical US city has incentive to use PT (if available). First, they find *C* = \$4 and *P* = \$2 per commute trip in a typical US urban area (2001 USD). From a regression analysis that uses data from 16 US cities they estimate *F* = 15 minutes. This means a rich commuter will prefer a car over PT at all *D* in the urban area if W_{Rich} > \$8 per hour (2001 USD) (i.e., the opportunity cost of time per minute for the rich commuter is at least \$0.13 or \$8 per hour). Conversely, the poor commuter will prefer PT at *D* close to the core (D ≈ 0) if W_{Poor} < \$8 per hour (2001 USD). Given the hourly wage for a poor worker in 2001, W_{Poor} < \$8 per hour (2001 USD) seems reasonable.

SI Text 3

All data and code mentioned below can be found in the zip file *DataandCode* under the banner "Working Papers, Data, and Code" at
https://www.bowdoin.edu/profiles/faculty/enelson2/index.html.

To estimate models (2.1) and (2.2) for each commuter group *j* run the Stata code *TechnologySpeeds.do* with the dataset *NHTS.dta*.

To replicate the multiple imputation analysis run the Stata code *ImputedMSACAT1.do*, *ImputedMSACAT2.do*, and *ImputedMSACAT3.do* with the datasets *MSACAT1Stata.dta*, *MSACAT2Stata.dta*, and *MSACAT3Stata.dta*.

To replicate the multinomial logit estimates of model (3.4) for each income class (poor or rich) – geography (1, 2, or entire urban area) – MSA category (1, 2, or 3) combination run the R scripts *MSACAT1PoorwithDistStandardize* and *MSACAT1RichwithDistStandardize* with the dataset *MSACAT1.csv*; *MSACAT2PoorwithDistStandardize* and *MSACAT2RichwithDistStandardize* with the dataset *MSACAT2.csv*; and *MSACAT3PoorwithDistStandardize* and *MSACAT3RichwithDistStandardize* with the dataset *MSACAT3.csv*.

To determine the *ceteris paribus* impact of a 1 standard deviation (SD) change in a mean commuter's independent variable on their mode choice probabilities (or in the case of a dummy variable, a change in the variable's binary status) according to estimated model (3.4) for each income class (poor or rich) – geography (1, 2, or entire urban area) – MSA category (1, 2, or 3) run the R scripts *MSACAT1PoorwithDistStandardize* and *MSACAT1RichwithDistStandardize* with the dataset *MSACAT1.csv*; *MSACAT2PoorwithDistStandardize* and *MSACAT2RichwithDistStandardize* with the dataset *MSACAT2.csv*; and *MSACAT3PoorwithDistStandardize* and *MSACAT3RichwithDistStandardize* with the dataset *MSACAT3.csv*.

To replicate the multinomial logit estimates of model (3.5) for each income class (poor or rich) – geography (1, 2, or entire urban area) – MSA category (1, 2, or 3) and imputation combination (recall there are 10 sets of *T* for each commuter) run the R scripts *MSACAT1PoorwithTimetoWkStandardize* and *MSACAT1RichwithTimetoWkStandardize* with the dataset *MSACAT1Imputed.csv*; *MSACAT2PoorwithTimetoWkStandardize* and *MSACAT2RichwithTimetoWkStandardize* with the dataset *MSACAT2Imputed.csv*; and *MSACAT3PoorwithTimetoWkStandardize* and *MSACAT3RichwithTimetoWkStandardize* with the dataset *MSACAT3Imputed.csv*.

To determine the *ceteris paribus* impact of a 1 SD change in a mean commuter's independent variable on their mode choice probabilities (or in the case of a dummy variable, a change in the variable's binary status) according to estimated model (3.5) for each income class (poor or rich) – geography (1, 2, or entire urban area) – MSA category (1, 2, or 3) and imputation combination (recall there are 10 sets of *T* for each commuter) run the R scripts *MSACAT1PoorwithTimetoWkStandardize* and *MSACAT1RichwithTimetoWkStandardize* with the dataset *MSACAT1Imputed.csv*; *MSACAT2PoorwithTimetoWkStandardize* and *MSACAT2RichwithTimetoWkStandardize* with the dataset *MSACAT2Imputed.csv*; and *MSACAT3PoorwithTimetoWkStandardize* and *MSACAT3RichwithTimetoWkStandardize* with the dataset *MSACAT3Imputed.csv*.

To replicate Table 8 (LA-2003) run the Stata do flies *LA2003Income.do* and *LA2003Poverty.do* with the data files *LA2003Income.dta* and *LA2003Poverty.dta*

To replicate Table 9 (LA-2005) run the Stata do files *LA2005Income.do* and *LA2005Poverty.do* with the data files *LA2005Income.dta* and *LA2005Poverty.dta*

To replicate Table 10 (Minneapolis-2004) run the Stata do files *Minneapolis2004Income.do* and *Minneapolis2004Poverty.do* with the data files *Minneapolis2004Income.dta* and *Minneapolis2004Poverty.dta*

To replicate Table 11 (Denver-1994) and Table 12 (Denver-2006) run the Stata do files *Denver19942006Income.do* and *Denver19942006Poverty.do* with the data files *Denver19942006Income.dta* and *Denver19942006Poverty.dta*

To replicate Table 13 (Phoenix-2008) run the Stata do files *Phoenix2008Income.do* and *Phoenix2008Poverty.do* with the data files *Phoenix2008Income.dta* and *Phoenix2008Poverty.dta*

SI Text 4

An analysis of 2017 US Census data (ACS 2017) sheds some more light on commute mode choices across the urban and HH income gradient. The 2017 survey includes people that could have responded to the questionnaire at any point between 2013 and 2017. Among many questions, a respondent is asked how much they earned in the past 12 months and their typical commute mode: 1) drove alone to work, 2) drove to work in a carpool, and 3) used public transit to commute to work. In the publicly available form of the ACS, the number of workers that used each form of commuting and the median annual income among each of these commuter groups is reported at the CT-level.

Compared to 2017 NHTS results, the 2017 ACS finds PT to be even less relevant to US commuters, even in MSA category 1 urban areas (SI Table 2). Whereas the 2017 NHTS estimated that over 50% of commuters used PT in the densest (1st ring) neighborhoods of the US largest cities with heavy rail, the ACS estimates only 4.31%. The mismatch between the NHTS and ACS results also hold, albeit to a lesser degree than in the 1st ring, in the 2nd ring and overall MSA category 1 urban areas. In MSA category 2 urban areas, however, the NHTS and ACS results match.

Like the 2017 NHTS, the 2017 ACS does suggest that richer commuters in MSA category 1 urban areas are much more able and/or willing to use PT than their counterparts in MSA category 2 urban areas. To get a better handle on the income distribution of PT versus car commuters in the ACS I plotted of the ratio of PT user median income to car (alone) commuter median income from every CT in MSA category 1 and 2 urban areas (SI Fig. 2). The plot indicates many CTs have ratios above 1. Interestingly, the CTs with the greatest ratios are found in the third rings of MSA category 1 and 2 urban areas (less than 10,000 people per square mile in category 1 and less than 2,000 people per square mile in category 2). In these areas the (unweighted) average and median ratio across all CTs is greater than 1 as well. I suspect this plot reveals that the few commuters that use PT in the outer ring tend to have very high paying jobs downtown (most suburban PT feeds workers to the downtown area). Given their high opportunity cost of time the ability to work while commuting is likely to be very appealing to such workers. Further, the less affluent from these areas are more likely to drive to work (and have no time for work while commuting) either because they are less time sensitive or have jobs outside of downtowns. This last dynamic means more car commuting because PT doesn't tend to connect suburbs to non-downtown work site. Further, this last dynamic tends to be concentrated among the less affluent because non-downtown jobs pay less, all else equal.

SI Text 5. Multiple imputation with predictive mean matching

According to the Stata 15.1 manual on Multiple imputation predictive mean matching works in the following manner. First, a linear model with the dependent variable of travel time for mode j and a covariate vector comprised of 2017 NHTS data is estimated using least squares across all commuters with observed travel times for mode j. Second, new model parameters are simulated from their joint posterior distribution under the conventional noninformative improper prior $Pr(\beta,\sigma^2) \propto 1/\sigma^2$ where β and σ^2 are the least square estimates of the mode j travel time model. Using the simulated model, mode j travel time is then predicted for all commuters that did not use this mode. Suppose commuter *i* is missing travel time for mode *j*. Suppose its predicted travel time using mode *j* is *X*. The 5 commuters that have measured travel time for mode *j* closest to *X* are matched to commuter *i*, mode *j*. In imputation iteration 1 one of the nearest 5 travel time measures is randomly assigned to commuter *i*, mode *j*, in imputation iteration 2 one of the nearest 5 lake's travel time measures is randomly assigned to commuter *i*, mode *j*, etc.

SI Text 6. Direction of omitted variable bias in estimates of the mode choice models (3.4) and (3.5) without the independent variable vehicle per household driver

In section 3 of the paper I find that a 1 standard deviation (SD) increase in income at the mean poor commuter HH's has little impact on the commuter's mode choice probabilities. However, one could push back on this claim with the reasonable sounding assumption that an increase in a poor commuter's HH income is likely to lead to the purchase of an additional car. Further, if increases in HH income are correlated with increases in cars at a HH then one could reasonably claim that most of the positive correlation between cars owned by a HH and the poor commuter's probability of picking the car mode can be traced back to an increase in HH income. To address this possibility, I re-estimated (2.4) and (2.5) without the independent variable vehicle per household driver (*V*). In this case the full impact of an increase in HH income on mode choice probabilities is not muted by the intermediate step using additional income to purchase a car.

I find that dropping *V* from mode choice models (3.4) and (3.5) means a 1 SD increase in the mean poor commuter's HH income increases their probability of choosing the car mode by an additional 5 percentage points or so compared to the probabilistic impact of increases in HH income when *V* is included in the model. However, this augmented impact of increases in HH income on car choice probabilities among poor HH commuters still does not match the impact of a 1 SD increase in *V* on poor commuter's car choice probabilities when HH income and *V* are both included in the mode choice model. Therefore, while poor HHs do tend to increase their car supply as they get richer and this increased supply leads to more car commuting, Glaeser et al.'s claim that a significant increase in income in a poor HH will lead a "massive shift from public transportation to driving" (p. 13) is unfounded.

And even the estimated additional 5 percentage points or so in the probability of using a car to commute I found after dropping *V* from the models may be an overestimate of the true effect. By dropping *V* from the models, I introduce omitted variable bias into the re-estimates of (3.4) and (3.5). Further, if the direction of this bias on HH income's car mode coefficient among poor HHs is positive then I am the overestimating the additional impact a increase in HH income has on the probability of a poor commuter switching to car use when *V* is not included in the mode choice models.

In the text we assume the true model has the form,

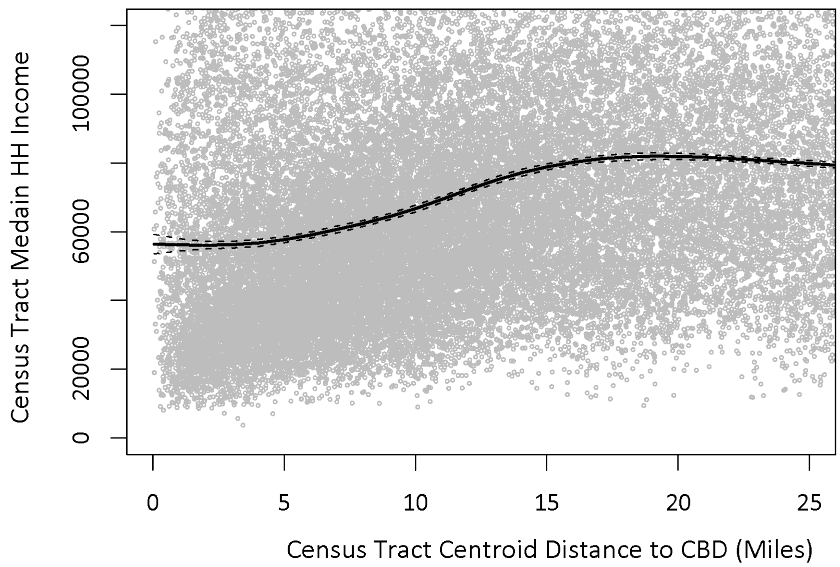
$$Mode\ choice_i = \alpha + \beta_1 I_i + \beta_2 V_i + \boldsymbol{\theta}\mathbf{X}_i + \epsilon_i \tag{A}$$

where *i* indexes commuters, *I* indicates annual HH income, and $\mathbf{X}_i$ contains the other covariates used in (3.4) or (3.5). To make an educated guess on the direction of bias in my model when *V* is omitted I use OLS to estimate the following equation for poor commuters from each location-geography combination,

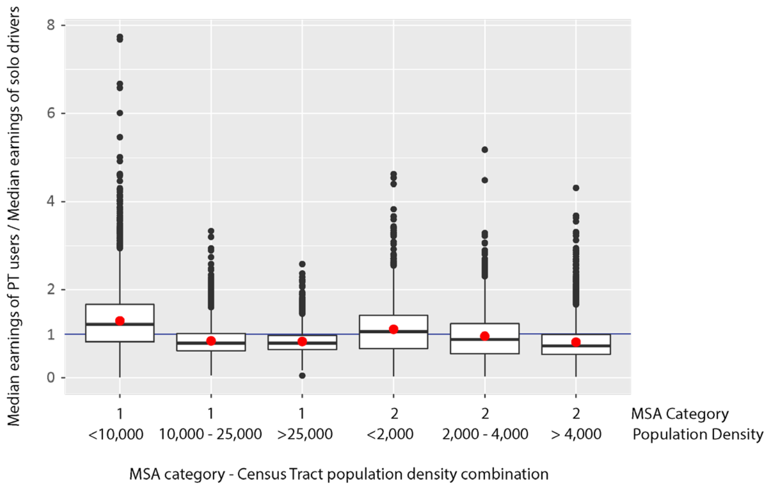
$$V_i = \alpha + \delta_1 I_i + \boldsymbol{\mu}\mathbf{X}_i + \varepsilon_i \tag{B}$$

where *i* indexes commuters and as in the models (3.4) and (3.5), all explanatory variables are standardized within each unique location-geography combination. Given that $\hat{\beta}_2$ from (A) is positive for poor commuters when the mode choice is car, a $\hat{\delta}_1 > 0$ (i.e., *V* and *I* are positively correlated) would suggest that the estimated coefficient on *I* in the poor commuter versions of (3.4) and (3.5) where *V* is dropped is *overestimated*. I say suggest because the final judgement on the direction of the bias is also a function of the correlations between *V* and the variables in **X**.

As can be seen in SI Table 5, $\hat{\delta}_1$ is consistently greater than 0. Therefore, this suggests that my finding that increases in HH income are responsible for an additional 5 percentage points in the probability that a poor commuter chooses the car mode when their HH gets 1 SD richer if we account for the impact of increasing income on car purchases is an *upper bound* on an increase in income's true effect on car commuting behavior.



SI Figure 1. Median HH income of every census tract (CT) found in all American CBSAs against each CT's Euclidean distance to its respective central business district (CBD). A spline fit to this data has a positive slope: expected HH income increases with distance from the CBD when you include data from every American CBSA. Glaeser et al. (2008) find that US census tracts (CTs) closer to CBDs are richer on average than those farther away as of 2000 across 16 US urban areas. These contradictory results suggest that the 16 US urban areas Glaeser et al. (2008) use for this analysis not representative of the US in general. Run the R script *FigureS1* to replicate this figure.



SI Figure 2: **The ratio of the median public transit (PT) commuter's 12-month earnings to the median solo car commuter's 12-month earnings in various MSA category – CT population density categories according to the 2017 ACS.** CTs in each MSA category are further categorized by their population density (people per square mile). The red circle is the mean ratio across the set of CTs in each unique combination of MSA category - CT population density set. The box indicates the range of CT ratios that fall in the 25th through 75th percentile of a MSA category – CT population density combination distribution of ratios. The black bar indicates the median ratio in each unique combination of MSA category - CT population density set. Each black point is the ratio in an outlier CTs.

SI Table 1. CBSA - MSA Crosswalk

		MSA Category			
CBSA ID	**CBSA**	**1**	**2**	**3**	**4**
12060	Atlanta-Sandy Springs-Marietta, GA	1	0	0	0
12580	Baltimore-Towson, MD	1	0	0	0
14460	Boston-Cambridge-Quincy, MA-NH	1	0	0	0
16980	Chicago-Joliet-Naperville, IL-IN-WI	1	0	0	0
17460	Cleveland-Elyria-Mentor, OH	1	0	0	0
31080	Los Angeles-Long Beach-Anaheim, CA	1	0	0	0
33100	Miami-Fort Lauderdale-Pompano Beach, FL	1	0	0	0
35620	New York-Northern New Jersey-Long Island, NY-NJ-PA	1	0	0	0
37980	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	1	0	0	0
39300	Providence-New Bedford-Fall River, RI-MA	1	0	0	0
40140	Riverside-San Bernardino-Ontario, CA	1	0	0	0
41860	San Francisco-Oakland-Fremont, CA	1	0	0	0
41940	San Jose-Sunnyvale-Santa Clara, CA	1	0	0	0
47900	Washington-Arlington-Alexandria, DC-VA-MD-WV	1	0	0	0
12420	Austin-Round Rock-San Marcos, TX	0	1	0	0
13820	Birmingham-Hoover, AL	0	1	0	0
15380	Buffalo-Niagara Falls, NY	0	1	0	0
16740	Charlotte-Gastonia-Rock Hill, NC-SC	0	1	0	0
17140	Cincinnati-Middletown, OH-KY-IN	0	1	0	0
18140	Columbus, OH	0	1	0	0
19100	Dallas-Fort Worth-Arlington, TX	0	1	0	0
19740	Denver-Aurora-Broomfield, CO	0	1	0	0
19820	Detroit-Warren-Livonia, MI	0	1	0	0
24340	Grand Rapids-Wyoming, MI	0	1	0	0
25540	Hartford-West Hartford-East Hartford, CT	0	1	0	0
26420	Houston-Sugar Land-Baytown, TX	0	1	0	0
26900	Indianapolis-Carmel, IN	0	1	0	0
27260	Jacksonville, FL	0	1	0	0
28140	Kansas City, MO-KS	0	1	0	0
29820	Las Vegas-Paradise, NV	0	1	0	0
31140	Louisville/Jefferson County, KY-IN	0	1	0	0
32820	Memphis, TN-MS-AR	0	1	0	0
33340	Milwaukee-Waukesha-West Allis, WI	0	1	0	0
33460	Minneapolis-St. Paul-Bloomington, MN-WI	0	1	0	0
34980	Nashville-Davidson-Murfreesboro-Franklin, TN	0	1	0	0
35380	New Orleans-Metairie-Kenner, LA	0	1	0	0
36420	Oklahoma City, OK	0	1	0	0
36740	Orlando-Kissimmee-Sanford, FL	0	1	0	0
38060	Phoenix-Mesa-Glendale, AZ	0	1	0	0
38300	Pittsburgh, PA	0	1	0	0
38900	Portland-Vancouver-Hillsboro, OR-WA	0	1	0	0
39580	Raleigh-Cary, NC	0	1	0	0
40060	Richmond, VA	0	1	0	0
40380	Rochester, NY	0	1	0	0
40900	Sacramento--Arden-Arcade--Roseville, CA	0	1	0	0
41180	St. Louis, MO-IL	0	1	0	0
41620	Salt Lake City, UT	0	1	0	0

		MSA Category			
CBSA ID	**CBSA**	**1**	**2**	**3**	**4**
41700	San Antonio-New Braunfels, TX	0	1	0	0
41740	San Diego-Carlsbad-San Marcos, CA	0	1	0	0
42660	Seattle-Tacoma-Bellevue, WA	0	1	0	0
45300	Tampa-St. Petersburg-Clearwater, FL	0	1	0	0
47260	Virginia Beach-Norfolk-Newport News, VA-NC	0	1	0	0
None	None	0	0	1	1

SI Table 2. Percentage breakdown of commute mode choices in 2017 ACS.

MSA category	Pop. density in respondents' home CTs	% of PT commuters (ACS)	% of PT commuters (NHTS)	% of CTs where med. inc. of PT commuters is greater than med. Inc. of drive alone commuters
1	All	4.80	17.83	29.25
	> 25,000	4.31	54.54	19.47
	10,000 – 25,000	4.53	23.14	20.15
2	All	4.85	5.05	9.67
	> 4,000	5.50	8.32	11.05
	2,000 – 4,000	4.59	3.35	10.64

Notes: "% of PT commuters (ACS)" is the weighted average of CT-level (HC04_EST_VC41 / (HC02_EST_VC41 + HC03_EST_VC41 + HC04_EST_VC41)) x 100 where HC02_EST_VC41 = "Car, truck, or van -- drove alone; Estimate; EARNINGS IN THE PAST 12 MONTHS (IN 2017 INFLATION-ADJUSTED DOLLARS) FOR WORKERS - Workers 16 years and over with earnings", HC03_EST_VC41 = "Car, truck, or van -- carpooled; Estimate; EARNINGS IN THE PAST 12 MONTHS (IN 2017 INFLATION-ADJUSTED DOLLARS) FOR WORKERS - Workers 16 years and over with earnings", and HC04_EST_VC41 = "Public transportation (excluding taxicab); Estimate; EARNINGS IN THE PAST 12 MONTHS (IN 2017 INFLATION-ADJUSTED DOLLARS) FOR WORKERS - Workers 16 years and over with earnings." The weight for each CT is given by (HC02_EST_VC41 + HC03_EST_VC41 + HC04_EST_VC41). "*Median earnings, drove alone* at the CT-level given by "Car, truck, or van - drove alone; Estimate; Median earnings (dollars)" (HC02_EST_VC51). *Median earnings, public transit* at the CT-level given by "Public transportation (excluding taxicab); Estimate; Median earnings (dollars)" (HC04_EST_VC51). Identification data that would let me map census tracts from the ACS to MSA category 3 areas in the NHTS does not exist.

SI Table 3

See the Excel workbook SITable3.xlsx for simulated changes in $\widehat{P}_{j,Inc,Geog,MSA}$.

SI Table 4

See the Excel workbook SITable4.xlsx for simulated changes in $\hat{p}_{j,Inc,Geog,MSA}$.

SI Table 5. OLS estimates of model (B) from SI Text 6. Standard errors are in parentheses.

	MSA Category 1	**MSA Category 2**	**MSA Category 3**
Geography	*All*		
	-0.23	-0.215	-0.051
(Intercept)	(0.052)	(0.042)	(0.032)
	0.186	0.179	0.165
HH Income	(0.031)	(0.024)	(0.017)
	0.051	0.054	0.138
Distance to work	(0.032)	(0.025)	(0.017)
	0.08	0.059	0.038
Age	(0.032)	(0.025)	(0.019)
	0.044	-0.147	-0.068
Male	(0.062)	(0.048)	(0.034)
	0.18	0.132	0.019
White	(0.063)	(0.048)	(0.035)
Geography	*1st ring*		
	-0.669	0.184	0.317
(Intercept)	(0.135)	(0.029)	(0.036)
	0.398	0.093	0.121
HH Income	(0.093)	(0.031)	(0.038)
	-0.033	0.093	-0.048
Distance to work	(0.094)	(0.032)	(0.042)
	-0.048	-0.087	-0.201
Age	(0.085)	(0.059)	(0.067)
	-0.026	0.187	0.145
Male	(0.179)	(0.059)	(0.068)
	-0.16	-0.311	-0.052
White	(0.192)	(0.082)	(0.065)
Geography	*2nd ring*		
	-0.218	-0.311	-0.052
(Intercept)	(0.097)	(0.082)	(0.065)
	0.149	0.224	0.067
HH Income	(0.065)	(0.055)	(0.032)
	0.044	0.098	0.102
Distance to work	(0.059)	(0.051)	(0.035)
	0.208	-0.11	-0.007
Age	(0.066)	(0.048)	(0.041)
	0.105	-0.051	0.109
Male	(0.113)	(0.095)	(0.07)
	0.238	0.306	-0.165
White	(0.117)	(0.098)	(0.071)