Journal of Urban Economics 108 (2018) 170–191

Contents lists available at ScienceDirect

Journal of Urban Economics

journal homepage: www.elsevier.com/locate/jue

Urban form and driving: Evidence from US cities☆

Gilles Duranton[a,*], Matthew A. Turner[b,c]

[a] Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104, United States
[b] Department of Economics, Box B, Brown University, Providence, RI 02912, United States
[c] IGC, NBER and PERC (Brown University), Providence, RI 02912, United States

A R T I C L E   I N F O

JEL classification:
L91
R41

Keywords:
Urban form
Vehicle-kilometers traveled
Congestion

A B S T R A C T

We estimate the effect of urban form on driving. We match the best available travel survey for the US to spatially disaggregated national maps that describe population density and demographics, sectoral employment and land cover, among other things. To address inference problems related to sorting and endogenous density, we develop an estimator that relies on an assumption of imperfect mobility and exploit quasi-random variation in subterranean geology. The data suggest that increases in density cause small decreases in individual driving. Applying our estimates to the observed distribution of density and driving in the US suggests that plausible densification policies cause decreases in aggregate driving that are small, both absolutely and relative to what might be expected from gas taxes or congestion charging.

1. Introduction

We estimate the effect on household driving of a variety of characteristics that describe urban form and cities, in particular the density of population and employment surrounding households' place of residence. In essence, we are asking how much less (or more) households drive as the area around their residence changes, in particular, as we make it denser.

We find that urban density has a small causal effect on individual driving. In most of our estimations 'urban density' is the density of residents and jobs within a 10-kilometer radius of where a driver lives. We find that the elasticity of vehicle kilometers traveled (VKT) with respect to this measure of density is between −7% and −10%. Put differently, a 10% increase in population density leads to a 0.7% to 1% decline in driving, all else equal. This result is not sensitive to the particular measure of density, but is sensitive to the scale at which we measure density. Residents and employment more than 10 km from a driver's residence do not have a measurable effect on driving behavior, nor do other measures of urban form other than density. Note also that this elasticity of −7 to −10% reflects the net of several effects: (i) higher density reduces trip distance by making each destination closer, (ii) greater proximity may elicit more trips, and (iii) higher density increases the unit cost of travel by increasing congestion.

Our first main contribution is to improve the causal identification of the relationship between urban form and driving. Our econometric framework derives from a simple model of travel behaviour. As density increases, household travel is subject to countervailing forces. A higher density makes destinations closer, which reduces household travel distance. In turn, closer destinations lead to more trips and thus an increase in household travel distance. We show that an implication of this model is that location and household specific unobservables may be correlated with density and driving.

We pursue a number of strategies to address the issue that households with particular preferences for driving may sort into areas of particular density. Among them, we develop a novel approach to the sorting problem for a cross-section of residents that follows from an intuitive definition of sorting and an assumption of imperfect residential mobility. We also use instrumental variables to address the problem of unobserved local factors correlated with density that may also affect driving behavior.

Our second main contribution is to combine the best available travel survey data for the US with spatially disaggregated national maps that

☆ Financial support from the Canadian Social Science and Humanities Research Council and from the Sustainable Prosperity Network is gratefully acknowledged by both authors. Turner gratefully acknowledges the financial support and hospitality of the Property and Environment Research Center. We are also grateful to many conference and seminar audiences; to the editor Nate Baum-Snow, two anonymous referees, Marlon Boarnet, and Henry Overman for helpful comments; and to Andy Foster for a particularly helpful suggestion about our model. This research could not have been conducted without the help of our research assistants, Tanner Regan, Nicolas Gendron-Carrier, Prottoy Akbar and Rebbeca Lindstrom.

* Corresponding author.
E-mail addresses: duranton@wharton.upenn.edu (G. Duranton), matthew_turner@brown.edu (M.A. Turner).
URL: https://www.real-estate.wharton.upenn.edu/profile/21470/ (G. Duranton), http://www.econ.brown.edu/fac/Matthew_Turner/ (M.A. Turner)

https://doi.org/10.1016/j.jue.2018.10.003
Received 15 March 2018; Received in revised form 26 October 2018
Available online 3 November 2018
0094-1190/© 2018 Elsevier Inc. All rights reserved.

describe, among other things, population density and demographics, sectoral employment, and land cover. Relative to past literature, which we discuss below, we thus use more exhaustive data and novel and extensive measures of urban form.

Our third main contribution is to provide some important insights regarding a variety of policy proposals. First, land use change is a widely proposed policy response to the problem of urban congestion. For example: in a State of Arizona Department of Transportation professional paper, Kuzmyak (2012) concludes that "greater adherence to smart growth principles of compact, mixed-land use,... , may result in important reductions in average trip lengths and VMT [vehicle-miles traveled] demand on local and regional roads" while the US Department of Transportation states that "[t]ransportation demand is reduced when residential and commercial uses are planned to be within close proximity to each other... ".[1]

Changes to urban planning also play a prominent role in policy discussions of carbon abatement. The Fourth Assessment Report of the IPCC discusses land use as a potential policy tool to reduce the demand for automobile travel (e.g., section 5.5.1.1 of Intergovernmental Panel on Climate Change, 2007), the more recent Fifth Assessment suggests that "[u]rban densification in the USA over about 50 years could reduce fuel use by 9–16%" (Table 8.3, Intergovernmental Panel on Climate Change, 2014), and California's Senate Bill 375 (September 7, 2006) asserts that "it will be necessary to achieve significant additional greenhouse gas reductions from changed land use patterns and improved transportation".

With an elasticity of driving with respect to density of −0.07, our results imply that achieving a 20% reduction in VKT would require a 25-fold increase in density for everyone. Such a policy, which would require shutting down most of the US territory to human activity, is arguably extreme and is not politically feasible. More realistic densification policies may instead reallocate population across areas. This will reduce driving in areas that become denser but it will also increase driving in areas with declining population. Our estimates suggest that, overall, such policies will cause only modest decreases in aggregate driving. A comparison of the effects of densification policies with what is known about the effects of gasoline taxes and congestion prices suggests that densification policies are unlikely to be a cost effective way to reduce aggregate driving to reduce traffic congestion or mitigate driving related carbon emissions.

Three strands of literature are relevant to our inquiry. The first is the large literature on the relationship between urban form and driving. The second investigates the relationship between the characteristics of a place and behavior. The third examines the extent to which unobserved attributes of places affect the way that cities develop in these places.

The relationship between urban form and driving has received much attention from the literature and is the subject of several surveys, including (Ewing and Cervero, 2001; Handy, 2005; Cao et al., 2009; Ewing and Cervero, 2010; Boarnet, 2011), and (Stevens, 2017). The primary focus of this literature is relationship between urban form and either total travel distance by households (e.g., Bento et al., 2005; Brownstone and Golob, 2009) or the journey to work (e.g., Gordon et al., 1989; Giuliano and Small, 1993; Glaeser and Kahn, 2004).[2]

Research typically revolves around estimating the effect on driving behavior of the 'three D's' of urban form proposed in Cervero and Kockelman (1997); 'Density', 'Diversity' and 'Design'. That is: the density of residents or employment; the diversity of activity, in particular the extent to which residential and other uses are mixed; and, usually, characteristics of various transportation networks. Our data will allow us to investigate two of these three, the density and diversity of neighborhood population and economic activity, and to touch on the third, the characteristics of the neighborhood street network.[3]

The possibility that an individual or household's location choice may depend on their predisposition to travel is widely recognized and Cao et al. (2009) survey the econometric techniques that have been applied to the problem. However, the literature has yet to identify a good source of random or quasi-random variation in neighborhood choice. To the extent that the literature implements instrumental variables estimations to deal with sorting, it relies on variables such as race or housing stock age that seem unlikely to satisfy the relevant exogeneity condition and are subject to the conceptual problem we describe in Section 3. Panel data sets are almost unknown and those that are available describe small areas and samples. In this light, the approach to the problem of sorting that we develop below is an advance. In addition, the possibility that the neighborhood characteristics of interest may be correlated with unobserved characteristics that affect driving, our 'endogenous density' problem, is usually ignored.[4] Our empirical strategy is also an advance in this respect.

2. A simple model of urban form and driving

To motivate our empirical analysis, we first present a simple model of equilibrium driving behavior. It focuses on how density affects key tradeoffs in travel decisions and illuminates the inference problems that our empirical investigation must overcome. Consistent with the regressions below, we focus on total travel distance by households. This is (arguably) the measure of travel that has received the most attention from both the academic literature and policy makers because it maps fairly directly into congestion, local pollution, and carbon emissions.

Consider a location with unit area and population density X. A resident with income W derives utility from the consumption of a continuum of differentiated varieties $Q(.)$ of measure N and the numéraire good C,

$$U = C + \theta\,\delta \left(\int_{i=1}^{N} Q(i)\mathrm{d}i \right)^{\rho}, \tag{1}$$

where θ is a resident-specific term, δ is a location-specific term, and $0 < \rho < 1$. To consume a differentiated variety, the resident must make a dedicated trip. The cost of a unit of variety i is $\tau D(i)$ where $D(i)$ is the travel distance to variety i and τ parameterizes the cost of travel.[5] We imagine that restaurants and movie theaters as well as local recreational amenities such as parks or museums would each constitute a 'variety' in this context.

Residence in a location requires the consumption of a unit of housing at price P_h. The budget constraint of a resident is thus $C + P_h + \int_{i=1}^{N} \tau D(i)Q(i)\mathrm{d}i = W$. To keep the problem tractable we assume that: (i) there are 'enough' varieties so that residents never consume the full set of available varieties, (ii) varieties can only be consumed in unit quantity $Q(i) = 1$, and (iii) varieties are symmetrically located around the resident so that $D(i) = D$ for all varieties i.[6] The budget constraint simplifies to $C + P_h + N\tau D = W$. Next, we can substitute this budget

[1] http://www.fhwa.dot.gov/planning/processes/land_use/land_use_tools/page02.cfm#toc380582783 , September 17, 2015.

[2] The literature has also investigated the relationship between urban form and other travel outcomes, including pedestrian trips and energy consumption (e.g., Brownstone and Golob, 2009; Glaeser and Kahn, 2008; Blaudin de Thé and Lafourcade, 2015).

[3] Subsequent literature has added more 'D's', that further local characteristics such as destination accessibility, distance to transit, and demand management, which we do not deal with here.

[4] Blaudin de Thé and Lafourcade (2015) in an exception.

[5] We impose an 'iceberg' (multiplicative) specification for travel costs to keep the consumer program tractable. This type of specification is extremely standard to model trade in goods (Head and Mayer, 2014). Its gravity implications also appear to describe commuting patterns extremely well (Ahlfeldt et al., 2015).

[6] Besides imposing convenient functional forms, our simple model also ignores many common features of travel such as the possibility of chaining trips. In addition, we do not explicitly deal with commutes and other work-related trips. Some of these complications are addressed in our regressions below. Our priority is to develop a tractable framework to underpin our regressions and to highlight the key econometric challenges that we face.

constraint into the utility function and simplify to obtain

$$U(N) = W - P_h + \theta\delta\, N^{\rho} - N\tau D\,. \tag{2}$$

Assuming income is high enough, the maximization of utility with respect to the number (mass) of varieties implies the following number of trips,

$$N = \left(\frac{\rho\theta\delta}{\tau D}\right)^{\frac{1}{1-\rho}}. \tag{3}$$

This expression indicates that residents take more trips if they have a greater taste for differentiated varieties, θ. For instance, some residents may enjoy dining out more than others. More generally, θ captures an individual resident's propensity to travel. The number of trips also increases with δ. For instance, a neighborhood near a nice beach may generate more trips than a neighborhood near a dirty beach. Our model can capture this by assigning one location a higher value of δ. Residents also make more trips when they are cheaper. This can occur because the cost of travel, τ, is lower or because trip distance, D, is shorter. In turn, differences in τ and D across locations arise as locations differ in how congested they are and in how compact they are. In addition the number of trips increases with ρ, which measures the (opposite of the) concavity of the utility function with respect to differentiated varieties.

In this respect, note that our model treats all trips as symmetrical. This assumption is extreme since leisure trips are arguably more discretionary than commutes. We could easily introduce compulsory commute trips by subtracting the cost of a commute trip, say τD, from the budget constraint above. This would leave the optimal number of discretionary trips in equation (3) unchanged. While this would not change our qualitative conclusions, introducing commutes more explicitly would lead to much more complicated functional forms below.

We are ultimately interested in how travel distance relates to density around a resident. Total travel distance by a resident is given by,

$$Y \equiv N\,D = \left(\frac{\rho\theta\delta}{\tau}\right)^{\frac{1}{1-\rho}}\left(\frac{1}{D}\right)^{\frac{\rho}{1-\rho}}, \tag{4}$$

where the last equality results from the use of equation (3). Like the number of trips, travel distance also increases with θ and δ and decreases with the unit cost of travel τ and trip distance D. The latter effect arises because the demand for trips is elastic with respect to trip distance.

Density at a location affects the demand for travel through a number of channels. A higher density reduces trip distance through greater accessibility. In turn, this reduces travel distance for a given number of trips but it also makes trips cheaper and thus elicits more trips. In addition, a higher density increases the unit cost of travel through more congestion. The net effect of improved accessibility and increased congestion on travel distance is ambiguous.

More specifically, to model the reduction in travel distance per trip that comes with greater population density, we assume

$$D = X^{-\zeta}\,, \tag{5}$$

where we refer to ζ as the accessibility elasticity.[7] We assume a power function for this relationship (and others below) to preserve analytical tractability. We show in Section 4 that the implied log linear relationship between travel distance and density fits the corresponding empirical relationship closely.

We expect congestion to depend on aggregate travel in the location. To capture this stylized fact in our model, suppose that travel costs are

$$\tau = \left(X\,\overline{Y}\right)^{\phi}, \tag{6}$$

where $\overline{Y}$ is mean travel distance and ϕ measures the elasticity of travel cost per unit with respect to aggregate travel, which we refer to as the congestion elasticity. Consider a location of unit size with parameter δ and a cumulative distribution of residents $F(\theta)$. Mean travel distance is then given by $\overline{Y} = \frac{1}{X}\int Y(\theta)\mathrm{d}F(\theta)$.

By construction, individuals do not account for their impact on τ. Therefore, equilibrium levels of driving will be greater than socially optimal levels. Even if exogenous changes in density reduce congestion and increase utility, they do not remove the need for congestion pricing. We return to this point below.

Our model describes only automobile travel and ignores the possibility that density might affect mode. This simplifying assumption is motivated by two features of our data. First, as we will see below, about 89% of all trips are made by a privately-owned vehicle. By excluding non-car travel, we only exclude a small share of trips. Second, even at high densities, mode choice is not very sensitive to density. As in our model, our US data suggest that the economically important margin of adjustment is the amount of driving, not substitution between driving and other modes.

After defining $\overline{\theta} = \left[\frac{1}{X}\int \theta^{1/(1-\rho)}\mathrm{d}F(\theta)\right]^{1-\rho}$, an index of the preferences of residents in a location, using the definition of $\overline{Y}$ above and inserting equations (5) and (6) into (4) implies

$$Y = \theta^{\frac{1}{1-\rho}}\left(\frac{\rho\delta}{\overline{\theta}^{\frac{\phi}{1-\rho}}}\right)^{\frac{1}{1-\rho+\phi}} X^{-\frac{\phi-\zeta\rho}{1-\rho+\phi}}, \tag{7}$$

after simplifications.

This expression shows that if the congestion elasticity, ϕ, is larger than the product of the accessibility elasticity, ζ, and the utility term, ρ, then travel distance decreases with population density. Two forces are at play. Travel distance increases with population density because of improved accessibility. This increase in travel distance also depends on how much the consumption of differentiated goods that require travel is valued in utility terms. At the same time, the cost of travelling also increases with density because of rising congestion. It is only when $\phi > \zeta\rho$ that travel distance declines with population density.

Expression (7) describes how total travel distance varies with population density. This our main counterfactual of interest in our estimations below. In our model, density affects distance traveled only through travel choices but does so through three margins: the distance to each destination, the number of destinations, and the cost of travel by unit distance. A richer model would also consider that density also affects distance traveled indirectly through income effects arising from higher wages (agglomeration economies) and higher housing costs (diseconomies of scale in construction and higher land costs).[8] While we return to these issues in our policy discussion at the end of this paper, we focus for now on the direct travel implications of density.

Substituting equations (3)–(7) into the utility function (2) leads to:

$$\begin{aligned} U &= W - P_h + (1-\rho)(\theta\,\delta)^{\frac{1}{1-\rho}}\left(\frac{\rho}{\tau D}\right)^{\frac{\rho}{1-\rho}} \\ &= W - P_h + (1-\rho)\rho^{\frac{\rho}{1-\rho+\phi}}\,\delta^{\frac{1+\phi}{1-\rho+\phi}}\left(\frac{\theta}{\overline{\theta}^{\frac{\phi\rho}{1-\rho+\phi}}}\right)^{\frac{1}{1-\rho}} X^{\frac{\rho}{1-\rho+\phi}(\zeta(1+\phi)-\phi)}. \end{aligned} \tag{8}$$

Although density affects utility only through travel behavior and we ignore the complications arising from the effects of density on the wage W and the cost of housing P_h, the outcome is nonetheless ambiguous. In equation (8), the coefficient of density, $\frac{\rho}{1-\rho+\phi}(\zeta(1+\phi)-\phi)$, is complicated because aggregate travel distance affects individual driving that occurs through congestion, as described in equation (6). However, the

[7] As accessibility improves residents face both more and closer options. Our formulation reflects this tradeoff, albeit in a simple, reduced-form manner. See Couture (2014) for micro-foundations.

[8] In practice we expect indirect effects to be small. With a typical elasticity of the wage with respect to density of 4% (Combes and Gobillon, 2015) and an elasticity of distance travelled with respect to income of about 0.25 as we estimate below, a 10% increase in density would imply a 0.1% increase in distance travelled, a small fraction of our preferred estimate. This is consistent with our empirical results below that show only small changes in the estimated density coefficient when we control for household and local income.

term in τD in the first line of equation (8) makes it clear that utility increases with density when it reduces trip distance, D, more than it increases unit travel cost, τ. For $\rho < 1$, utility increases with density when the accessibility elasticity is large enough, $\zeta > \frac{\phi}{1+\phi}$.

Next, we note that it is only when the exponent on X is positive in equation (8) and negative in equation (7) that travel declines while utility increases with density. These two conditions require $\frac{\phi}{\rho} > \zeta > \frac{\phi}{1+\phi}$. That is, the accessibility elasticity, ζ, must be large enough that utility increases with density but not so large that travel also increases. It is only when parameters satisfy these particular conditions that the model both predicts a widely conjectured empirical relationship and satisfies a necessary condition to rationalize policies to increase population density.

It is also easy to see from equation (8) that $\frac{\partial^2 U}{\partial\theta\,\partial X} \geq 0$ when $\frac{\partial U}{\partial X} \geq 0$. In words, there is a positive complementarity between the propensity to take trips and population density when utility increases with population density. This complementarity occurs because, when trips become more valuable when density increases, they become all the more valuable to households who enjoy taking trips more.

In Section 6, we extend our model to solve for the location choices of residents. In this extension, we show that the single-crossing condition implied by this complementarity between the propensity to take trips and density leads to the perfect sorting of residents across locations of different density. More specifically, residents with a greater propensity to make trips choose to locate in denser locations. The opposite form of sorting occurs when utility decreases with density. Hence, in general, we expect a non-zero correlation between the propensity to make trips, θ, and population density, X, to be a feature of our data.[9] Importantly, the direction of the bias is ambiguous. When increases in population density lead to large improvements in accessibility, we expect residents with a higher propensity to travel to locate where density is higher. When increases in population density lead instead to small improvements in accessibility, we expect on the contrary residents with a higher propensity to travel to locate where density is lower. Hence, an OLS regression of distance travelled on population density may understate or overstate the true effect of density because of the sorting of residents.

In addition, it is also easy to see that, in general, $\frac{\partial^2 U}{\partial\delta\,\partial X} \neq 0$. Hence, as we allow households to choose where to reside, we should also expect a non-zero correlation between how beneficial trips are in a location, δ, and population density, X.

If residential sorting is perfect in equilibrium, then we must have $\overline{\theta} = \theta$. In fact, we expect sorting to be less precise than this, and our econometric model relies on the fact that residential mobility is imperfect. To describe such a process parsimoniously, we instead suppose that $\theta = \overline{\theta}\nu$, where ν is an error term. Using this relationship in equation (7) and taking logs then gives,

$$y = \frac{\log\rho}{1-\rho+\phi} - \frac{\phi-\zeta\rho}{1-\rho+\phi}\,x + \epsilon\,, \tag{9}$$

where

$$\epsilon = \frac{\log\delta}{1-\rho+\phi} + \frac{1}{1-\rho+\phi}\log\theta + \frac{\phi}{(1-\rho+\phi)(1-\rho)}\log\nu\,. \tag{10}$$

and $y \equiv \log Y$ and $x \equiv \log X$.

Equation (9) describes a regression of driving on density. This regression, typically conducted with cross-sectional survey data, forms the basis of the large literature described above. Because local gains from trips, δ, and the propensity to make trips, θ, are not observed, they enter the error term. Given their expected correlation with population density, the estimated coefficient of x is potentially biased. The sorting of travellers and the endogeneity of density are the two main identification challenges we face in our empirical work below.

3. Econometric model

We would like to estimate the relationship between urban form and driving behavior. We begin by considering the problem of sorting and then turn to the problem of endogenous urban form.

Each person (household) is assigned to a geographic unit. As we discuss below, these will be regular grid cells of approximately one kilometer square. For each such unit we construct measures of urban form, usually a measure of density, which we also discuss below. Let i index individuals and j index residential locations. We are interested in explaining how driving behavior y_{ij} varies with urban form. More specifically, we are interested in knowing how the driving behavior of a randomly selected person or household changes when we change urban form in or around their residential location.

Let x_j^0 denote the urban form variable of interest for geographic unit j at an initial period (density in the model above and much of empirical work below), usually around 1990 and let x_j^1 denote the urban form variable of interest usually around 2010, contemporaneous to y. Define $\Delta x_j = x_j^1 - x_j^0$. We observe both contemporaneous and historical descriptions of urban form at each location, but we observe each driver only once.

Suppose that driving for each person is described by the following equation,

$$y_{ij} = \theta_{ij} + \beta x_j + \delta_j, \tag{11}$$

so that observed driving for each person is determined by an individual specific intercept, θ_{ij}, a location specific intercept, δ_j, and the urban form in person i's location j, x_j. The parameter of interest, β, measures the effect of local urban form on distance travelled.

We note that this is equivalent to the equilibrium driving equation (9) derived above, where, in a slight abuse of notation, we renormalize θ_{ij} and δ_j to improve legibility. Importantly, in both equations (9) and (11) individual taste parameters and location specific effects enter only through the intercept. They do not lead to individual or neighborhood level differences in β, the rate at which individuals change their behavior in response to density. This simplifies our econometric task considerably and we appeal to the theoretical analysis above to justify this restriction. This assumption also finds some empirical support in our results: we perform our main regression on many different subsamples and do not find measurable differences in β across samples.

Given equation (11), our two main inference problems are that people do not choose their locations at random and that observed and unobserved attributes of urban form are correlated with, and may affect driving. We address each problem in turn.

To begin, suppose that individual specific intercepts are not observed, but are drawn from the real interval Θ, let w denote observable individual characteristics related to location choice and let the distribution of individual types at each location j be determined by

$$\theta_{ij} = \alpha_0 + \alpha_1 x_j + \alpha_2 w_{ij} + \mu_{ij}, \tag{12}$$

where μ is a random variable and $E(x_j\mu_{ij}) = 0$. That is, the assignment of types to location j depends on urban form, on observable individual characteristics, and on unobserved individual characteristics. If $\alpha_1 > 0$, then drivers with a larger θ sort into neighborhoods with a larger x and conversely. As μ increases, residents derive more utility from trips for reasons unrelated to x.

Using both equation (12) and (11), we have that

$$\begin{aligned} y_{ij} &= (\alpha_0 + \alpha_1 x_j + \alpha_2 w_{ij} + \mu_{ij}) + \beta x_j + \delta_j \\ &= \alpha_0 + (\alpha_1 + \beta)x_j + \alpha_2 w_{ij} + \epsilon_{ij}, \end{aligned} \tag{13}$$

[9] Aside from the direct channel based on travel preferences that we emphasize here, other forms of residential sorting could take place, including income sorting as mentioned above. We worry about these alternative forms of sorting only to the extent that they affect travel behavior. Our empirical approach to residential sorting does not rely on a particular channel.

where $\epsilon = \mu + \delta$. Thus, if $\alpha_1 \neq 0$ or $E(\epsilon_j x_j) \neq 0$, OLS estimates of β will be biased.

Our approach to this sorting problem relies on an assumption of imperfect mobility. We now consider two time periods $t = 0$ and $t = 1$ and suppose that at $t = 0$ all agents match to locations as described above. At $t = 1$ a randomly selected fraction, s_j, of these residents relocates and is replaced by agents who sort on the basis of current conditions. With these assumptions in place, for a location where $x_j^1 = x_j^0 + \Delta x_j$, expected driving at $t = 1$ is

$$\begin{aligned} y_{ij}^1 &= (1 - s_j)\Big[(\alpha_0 + \alpha_1 x_j^0 + \alpha_2 w_{ij} + \mu_{ij}) + \beta x_j^1 + \delta_j\Big] \\ &\quad + s_j\Big[(\alpha_0 + \alpha_1 x_j^1 + \alpha_2 w_{ij} + \mu_{ij}) + \beta x_j^1 + \delta_j\Big] \\ &= \alpha_0 + (\alpha_1 + \beta)x_j^0 + \alpha_1 s_j \Delta x_j + \beta \Delta x_j + \alpha_2 w_{ij} + \epsilon_{ij} \\ &= A_0 + A_1 x_j^0 + A_2 s_j \Delta x_j + A_3 \Delta x_j + \alpha_2 w_{ij} + \epsilon_{ij}. \end{aligned} \tag{14}$$

In fact, we will not always observe s_j directly. Instead, we observe characteristics that vary systematically with the mobility rate, e.g., driver age or mean housing tenure in the driver's home cell. To understand how this allows similar tests, denote our mobility proxy by $\tilde{s}$ and suppose that mobility varies with $\tilde{s}$ according to $s = g(\tilde{s})$. Taking a linear approximation, we have $s = \gamma_1 \tilde{s}$, where $\gamma_1 \neq 0$ is assumed. Substituting this expression for s into equation (14) we see that the coefficient on $\tilde{s}\Delta x$ is $\alpha\gamma_1$. Substituting into equation (14) gives

$$y_{ij}^1 = A_0 + A_1 x_j^0 + A_2 \gamma_1 \tilde{s}_j \Delta x_j + A_3 \Delta x_j + A_4 w_{ij} + \epsilon_j. \tag{15}$$

Equation (15) suggests two parametric tests of the importance of sorting. First, the difference between the coefficients of x^0 and Δx is α_1. This is the parameter that describes how the unobserved individual propensity to drive varies with urban form in equation (12). Since $\alpha_1 = A_1 - A_3$, we can reject the hypothesis that $\alpha_1 = 0$ by rejecting the hypothesis that $A_1 = A_3$. Second, we can reject the hypothesis that $\alpha_1 = 0$ by rejecting the hypothesis that $A_2\gamma_1 = 0$. In fact, our estimates will generally indicate the $A_2\gamma_1$ is tiny and not significantly different from zero. However, because this test compounds two structural coefficients, we regard it as less informative than tests based on the difference $A_1 - A_3$. Although they are imprecise, point estimates in our preferred specification suggest that $\alpha_1 < 0$ and is about one sixth the magnitude of β. That is, individuals with smaller propensity to drive move to dense places, but this sorting most likely makes only a modest contribution to the observed relationship between urban form and driving.

This methodology requires two comments. Identification rests on the assumption that as urban form changes, so do the characteristics of the marginal resident. Not only does this seem like a reasonable hypothesis, it also follows a common sense definition of 'sorting'. While we express the intuition precisely and in particular functional forms, the underlying intuition seems unrestrictive. Second, as we have described it, sorting affects only residents moving to a location, not those moving away from it. More realistically, we might expect a non-random sample of people to move from a location, and in the case of an increase in density, they should value density less highly than the average current resident, who in turn should value density less highly the average arrival. We generalize our framework to describe this intuition precisely in Appendix A. This leads to a similar empirical strategy.

While the estimation described in equation (15) addresses the problem of sorting by unobserved individual characteristics, it does not address the possibility of omitted location variables correlated with urban form or changes in urban form. For example, better natural amenities may lead to a greater concentration of residents as well as to more driving to enjoy them (or go around them). To address this problem, we consider the system of equations,

$$y_{ij} = \theta_i + \beta x_j + \delta_j\,, \tag{16}$$

$$x_j = \gamma_0 + \gamma_1 z_j + \eta_j\,. \tag{17}$$

In the context of this system, our omitted variables problem may be stated as $E(x_j\delta_j) \neq 0$. We resolve this problem by relying on instrumental variables estimation. As the system above suggests, this requires an instrument that predicts urban form but that does not otherwise affect driving, or more formally, that $\gamma_1 \neq 0$ and $E(z_j\delta_j) = 0$. In our empirical work, we rely on various measures of subterranean geology as instrumental variables. As we will see, these measures are important determinants of urban form and it is difficult to imagine other channels through which they could affect driving behavior than by affecting the urban form.

Although this is a standard instrumental variables estimation, in our context, it requires two comments. First, we should not expect our instrumental variables estimation to resolve the problem of sorting. To see this, let $\hat{x}_j = \gamma_0 + \gamma_1 z_j$ and rewrite equation (13) using (17) as,

$$\begin{aligned} y_{ij} &= \alpha_0 + (\alpha_1 + \beta)(\hat{x}_j + \eta_j) + \epsilon_j\,, \\ &= \alpha_0 + (\alpha_1 + \beta)\hat{x}_j + ((\alpha_1 + \beta)\eta_j) + \epsilon_j)\,. \end{aligned}$$

That is, as long as residents sort on the component of the urban form predicted by underground geology in the same way as they sort on the residual component, the instrumental variables regression does not lead to unbiased estimates of β. Thus, instrumental variables estimation can solve the problem of unobserved local characteristics, but it cannot solve the problem of unobserved individual characteristics.

In light of the intuition above, we would ideally implement our instrumental variables strategy in the context of equation (15) which explicitly accounts for sorting. In practice, our instruments are not able to predict changes in urban form, only levels. Thus, in spite of its theoretical appeal, this strategy is beyond the reach of our data. With this said, the data suggest that neither sorting nor omitted variables cause economically important biases in our estimates, so we can reasonably conjecture that allowing these two biases to interact would also be unimportant.

4. Data

Our analysis requires three main types of data; household and individual level travel behavior, a description of urban form for each household, and finally, a description of subterranean geology. To implement our response to the sorting problem, we require panel data describing urban form, but only cross-sectional travel data.

We also require a way of matching survey respondents to measures of urban form. To accomplish this, we construct a regular grid of 990-meter square cells by aggregating the 30-meter square cells that describe land cover. Each household is matched to the cell which contains the centroid of the household's census block group. We will refer to this cell as an individual or household's 'home cell', and in a slight abuse of language, describe cells as having an area of one square kilometer.[10] We convert all data describing urban form to this resolution as described below. With this data structure in place, we can construct urban form measures for each household on the basis of arbitrary geographies by averaging over the relevant sets of grid cells. In particular, we can examine the square kilometer surrounding each household by reporting the characteristics of its home cell, we can average over all cells within 10 km of the home cell or over all cells lying in the same MSA.

Data on individual travel behavior come from the 2008–2009 National Household Transportation Surveys (NHTS).[11] The NHTS reports

[10] Our data are projected onto a flat surface using an Albers Equal Area projection. This projection transforms our approximately round planet into a plane and preserves area by compressing the North-South dimension of pixels away from the equator. This preserves pixel area at the expense of pairwise pixel distances. As a practical matter, over the range of distances we consider, i.e. about 10 km, such cartographic details are not important.

[11] U.S. Department of Transportation, Federal Highway Administration (2009).

several measures of total annual driving for each household or individual in a nationally representative sample of households. Our main dependent variable is household annual vehicle kilometers travelled (VKT) and is reported in the first row of Table 1.[12] This measure of household annual mileage is computed by the survey administrators, "bestmiles', and is their preferred measure. In robustness checks, we consider four other measures of individual and household driving distance, stated annual vehicle kilometers traveled, a reported odometer measure of kilometers traveled, individual daily kilometers traveled on the survey day, and distance to work.

Table 1 reports descriptive statistics for several measures of driving from the NHTS. The three measures of total household driving have sample means of 37,022, 33,014 and 33,123 km over slightly different samples of households. Except where noted otherwise, we restrict attention to households and individuals who live in MSAs.[13] Aggregating individual VKT and travel time at the household level implies that households travel 73.2 km in 98.7 minutes at an average speed of 42.6 km per hour.[14] Individual distance to work is 22.5 km. These values reflect the sample of household members who filled out a travel diary reporting positive travel and those who reported driving to work. We also note that, on average, households conduct 89% of their trips with a privately-owned vehicle. The transit share represents less than 2%.[15]

The NHTS survey reports household and individual demographics. These demographic variables provide a description of household race, size, income, educational attainment, and home ownership status. Mean household income is $71,257 and the average over households of the average age of household adults is 53.5 years. We also note that nearly 90% of households in our sample are homeowners.

Urban form data are more complicated. To measure the share of developed land cover, we rely on the 1992, 2002 and 2006 National Land Cover Data (NLCD).[16] While the NLCD reports many land cover classifications, we sum the urban classes in each year to measure the share of urban cover in each grid cell. Table 1 reports descriptive statistics for our sample. For an average survey respondent, 4.40% of the land area within 10 km of their home cell is in urban cover in 2006.[17]

To assign 2000 census data to our grid cells, we distribute block group data to our grid cells using an area weighting based on a geocoded map of 2000 census block groups. We perform a similar exercise for 1990 and 2010.[18] With this correspondence between block groups and grid cells in place, we are able to assign any block group variable reported in the 1990, 2000 or 2010 census and in the American Community Survey (ACS) to our grid.[19] All urban form variables involving demographic characteristics are computed on this basis. Table 1 reports that for an average survey respondent, the average residential density within a radial distance of 10 km of their home cell is 755 per square kilometer.

Using ACS and census tabulations, we also measure a number of other local characteristics such as an average length of tenure of 10.3 years and a renter share of 26.0%. We use these variables in estimations below, and note that there is some variation across households in the mobility and tenure rates of their neighborhoods.

Employment data are based on zipcode business patterns. These data report both aggregate and sectoral employment by zipcode. We assign these data to our grid on the basis of zipcode maps using the same procedure that we use for census data.[20] We use zipcode business patterns for the years closest to the NHTS survey years, and to reduce measurement error, average over the nominal year of the survey and the preceding year.

For some of our results, we rely on the 2007 National Highway Planning Network map (Federal Highway Administration, 2005) to describe the road network. This map is part of the federal government's efforts to track roads that it helps to maintain or build. It describes all interstate highways and most state highways and arterial roads in urbanized areas. To construct measures of road density, for each grid cell containing a survey respondent, we construct disks of radius 5, 10 and 25 kilometers centered on this cell. For each such disk, we then calculate kilometers of each type of road network in that disk. In addition to these data we also use the PRISM gridded climate data (PRISM Climate Group at Oregon State University, 2012b; 2012a) to measure temperature and precipitation in each grid cell.

For much of our analysis, we use the total number of people living or working within 10 km of each survey respondent to measure urban form and call this measure "10-kilometer density'.[21] We sometimes also work with the corresponding measure based only on the household's home cell and call this measure '1-kilometer density'. When the scale of analysis is clear, we sometimes refer to these quantities as 'density'. Table 1 reports that for an average household survey respondent, the 10-kilometer density is 1072 per square kilometer. People and jobs tend to be denser nearer survey respondents" homes, 1-kilometer density is 1513.

Fig. 1 presents two probability distribution functions, the fine dashed black line for NHTS sample population and the heavy gray line for census population. Both distributions have a mode around 8 which, converting from logs to levels, corresponds to a density of about 3000 per square kilometer. While the two distributions of census and NHTS people are generally close, they diverge slightly at high densities. This confirms the slightly higher response rates of the NHTS in less dense locations (U.S. Department of Transportation, Federal Highway Administration, 2009).

Panel (a) of Fig. 2 illustrates the way that people in the US are exposed to our measure of 1-kilometer density. In this map, the white area contains the 10% of the US population living at the lowest density. This

12 Our initial NHTS sample contains 150,147 households of whom we can locate 149,638 on our grid. We have a positive measure of vehicle kilometers traveled for 136,530 households. After restricting our sample to those observations for which we have a full set of household and individual characteristics, we are left with 126,203 households, 99,875 of whom live in an MSA as defined in 1999.

13 This is purely for expositional convenience. It allows us to include MSA indicator variables in our regressions without changing our sample.

14 This is an average across households. Dividing aggregate VKT by aggregate travel time implies a speed of 44.5 km per hour. Couture et al. (2018) report a mean speed per trip of 38.5 km per hour. The differences between those numbers are due to the fact that shorter trips are slower. Averaging across trip gives them a greater weight than averaging total travel across households. In turn, a household average will also weight shorter trips more does the ratio of aggregate distance to aggregate travel time.

15 Walking represents 8.4% of all trips but only 4.3% of trips longer than one kilometer and less than 0.1% of household VKT. Biking trips and taxi trips are each less than 1% of trips.

16 United States Geological Survey (2000, 2011b) and United States Geological Survey (2011a).

17 Note that all densities for rings around a survey respondent's home are normalized by the number of grid cells for which we have population and employment information. This prevents us from underestimating density for households who live by the sea, a lake, or uninhabitable terrain.

18 The particular census maps we use are: Environmental Systems Research Institute (1998b, 2004); U.S. Department of Commerce, U.S. Census Bureau, Geography Division (2010).

19 Sources for these data are: Missouri Census Data Center (1990, 2000, 2010) and National Historical Geographic Information System (2010).

20 Sources for our zipcode maps are: Environmental Systems Research Institute (1998a); U.S. Department of Commerce, U.S. Census Bureau, Geography Division (2010).

21 In principle, a pixel with many employees and few residents may affect driving behavior of residents differently than does a pixel with the opposite ratio but same total, so that our aggregated density measure may introduce measurement error. In Table 9 we decompose density into its two components and see that the regression R^2 is unchanged, so that, in fact, the more aggregated density measure does not have less ability to explain driving. This suggests that inference problems introduced by our aggregate measure of density are probably not important.

Table 1
Descriptive statistics for NHTS households, MSA sample.

Variable	Mean	Std. Dev.	5th percentile	95th percentile	Observations
Vehicles km travelled (VKT)	37,022	29,826	4459	87,906	99,875
log VKT	10.17	1.01	8.40	11.38	99,875
Annual VKT	33,014	29,766	3645	82,620	93,602
Odometer VKT	33,123	24,647	6388	74,483	71,742
Household daily VKT	73.2	66.8	6.5	208.1	83,313
Household daily travel minutes	98.7	70.0	17	234	83,313
Household daily speed	42.6	38.8	13.9	75.6	83,313
Share of trips by POV	0.889	0.456	1	1	93,198
Distance to work	22.5	34.4	1.6	61.6	95,532
10-km density	1072	1559	44.9	3,222	99,875
log 10-km density	6.30	1.31	3.81	8.08	99,875
10-km population density	755	1027	34.7	2211	99,875
10-km share developed (%)	4.40	5.61	0.07	15.5	99,875

Notes: Authors' calculations for 2008–2009 (NHTS variables in rows 1–9), 2010 (census variables in rows 10–12), and 2006–2011(NLCD in row 13). Distances are measured in kilometers and monetary values in current American dollars. Household age is mean age for the adult members of the household. Household daily VKT, travel time, and speed are computed for all households with positive travel by summing all trips across the surveyed members of the household. Household speed is computed by dividing VKT by travel time for each household and averaging across households. 'Density' refers to the sum of jobs and residents unless it is qualified by employment or residential population. All densities are reported per square kilometer. POV refers to privately-owned vehicles.

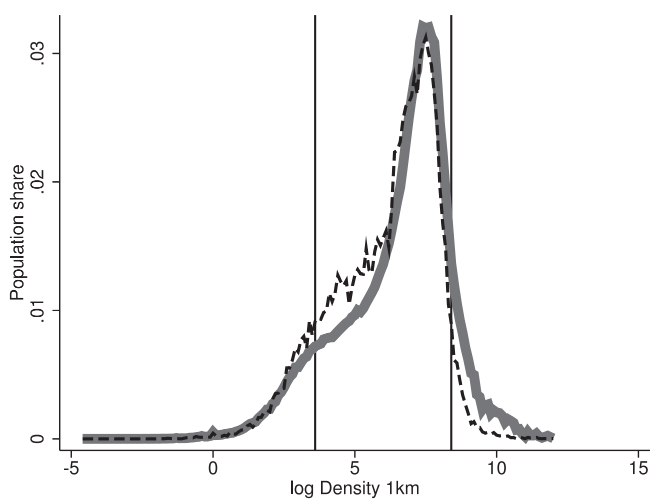

Fig. 1. The distribution of population conditional on density.
Notes: The dashed black line describes the distribution of people surveyed by the NHTS. We calculate number of NHTS people in each cell and then take the share of the total NHTS population living in cells of given densities and represent this on a log scale. The heavy gray line provides the corresponding information for census population, i.e., for the whole contiguous continental US. These distributions are based on the whole sample of the NHTS for which we record household VKT, not the MSA only sample on which we base most of our regressions and Table 1. Dropping the non-msa observations to be consistent with our regressions only affects the lower tail of the distribution. The two vertical lines indicate bottom and top density deciles.

region is about 5.8 million square kilometers and 83% of the land area of the continental US. On average, the about 30 million people living in this region have 6.25 people or jobs in their home cell. The barely visible black areas in this map contain the 10% of the US population living at the highest densities. This area is less than 15,000 square kilometers and about 0.2% of the land area of the continental US. On average, residents of these areas share their home cells with about 5421 other people and workers. That is, the decile of US population living at the highest densities lives at densities about 870 times higher than the lowest density decile. The medium gray area in this figure houses the residual 80% of the population.

In our instrumental variables estimation, we rely on variables constructed from United States Geological Survey ([2001,2003,2005]). United States Geological Survey (2003) describes the incidence of aquifers in the continental US. Using this map, we determine which grid cells overlay consolidated or semi-consolidated aquifers. Panel (b) of Fig. 2 illustrates these pixels. Burchfield et al. (2006) find that an MSA level index of aquifer prevalence is a good predictor of an aggregate measure of urban form. We will also find aquifers are good predictors of local density. Usefully, the map indicates that aquifers are broadly distributed across the country so that instrumental-variables estimates will not be driven by variation within particular small regions. United States Geological Survey (2005) describes a measure of earthquake intensity that ranges from 0 to 18. We consolidate to three categories; low, medium or high earthquake exposure. Panel (c) of Fig. 2 illustrates these regions. Areas of high earthquake intensity are dark. United States Geological Survey (2001) describes landslide susceptibility. The source data contains six categories, which we consolidate to low, medium and high risk. Panel (d) of Fig. 2 illustrates high risk areas in dark gray, medium risk areas in light gray and low risk areas in white. Like the aquifers map, neither landslide nor earthquake risk are concentrated in small geographic areas so that instrumental variables estimates based on these variables are not driven by small regions of the country.

Bringing together our data about travel behavior and urban form, Fig. 3 describes mean per person driving as a function of density. Except for extremely high or extremely low levels of density, the logarithmic scale of the figure shows a clear log linear trend. On this basis, we rely on multiplicative rather than additive regression equations. The far left part of the figure is noisy but this occurs for levels of log density below 2 and concerns less than 1% of the observations for our preferred sample of households that live in MSAs. The far right part of the figure is also noisy but, again, less than 1% of our observations live at a log density above 9. Indeed, only about 260 cells have a log density above 10 and of these, about two thirds are in the New York MSA. Despite this noise, the figure suggests that the effect of density on driving may increase at very high levels of density. We will look for this sort of non-linearity in our regressions but recall that these areas represent a tiny part of the country that may differ from the rest in many ways other than density.

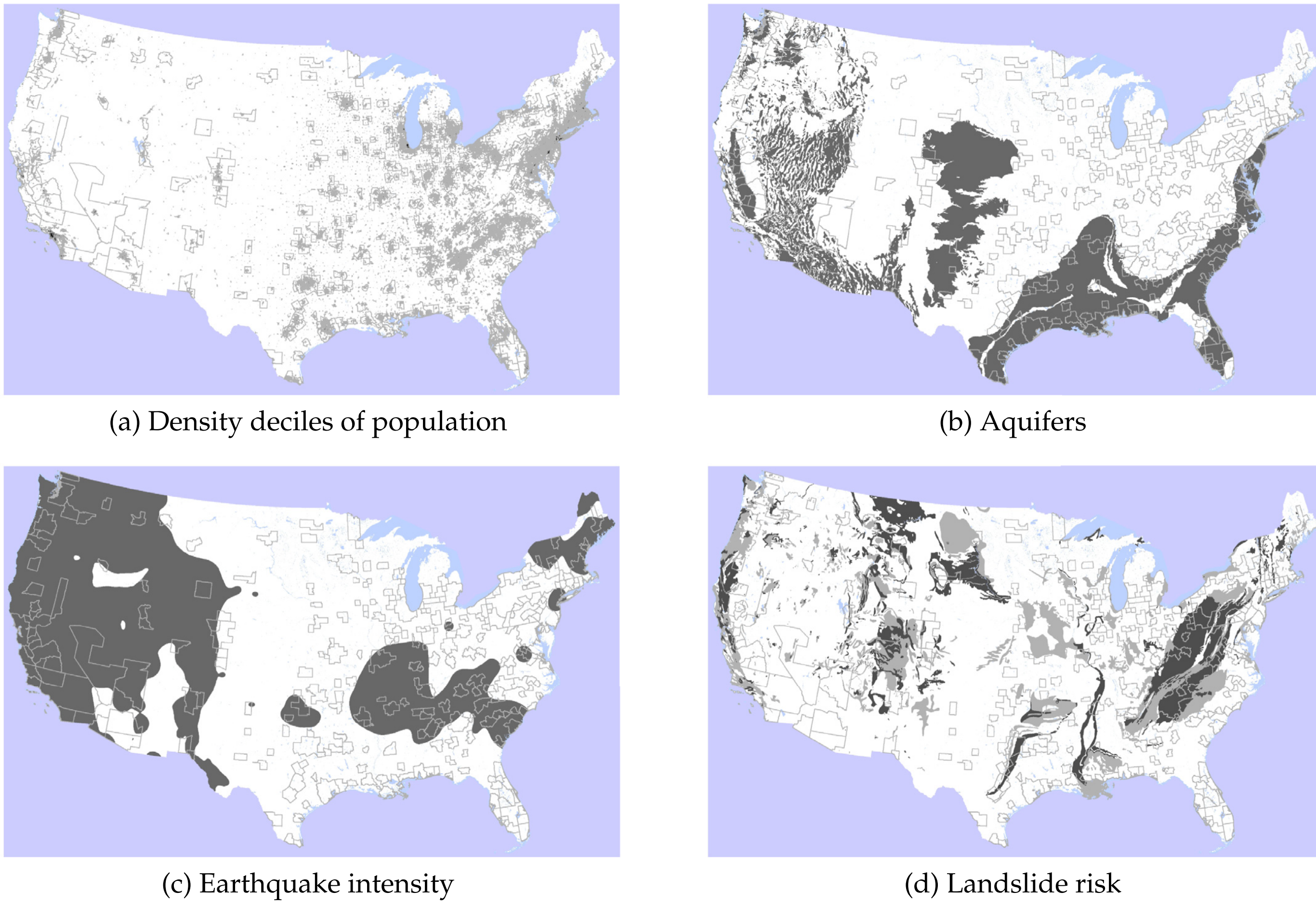

Fig. 2. Maps.
Notes: Panel(a): White indicates the area inhabited by people living in the bottom decile of density. Black indicates the area inhabited by people living in the top decile of density. Gray indicates the area inhabited by the 80% of the people living at intermediate densities. Panel (b): Gray indicates areas overlying unconsolidated or semi-consolidated aquifers and white indicates the absence of such aquifers. Panel (c): Darker gray indicates areas subject to larger earthquakes. Panel (d): Darker gray indicates areas subject to higher landslide risk. 2000 MSA boundaries shown in light gray in all four maps.

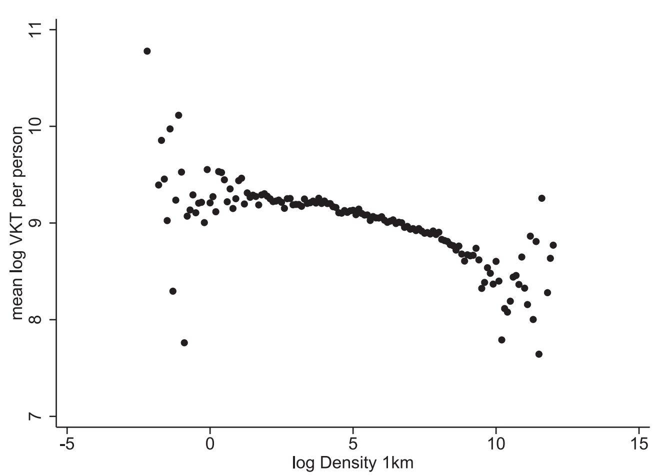

Fig. 3. Vehicle-kilometers traveled and density.
Notes: We first calculate mean per person VKT for each home cell by dividing household driving by the count of household members. We then calculate mean VKT conditional on density as density varies. Both axes are in log scale.

5. Results

We proceed in steps. First, we present OLS results showing the relationship between our preferred measures of driving and urban form, household VKT and the density of residents and jobs within 10 km. Second, we verify that these relationships are robust to different measures of driving and to the scale at which we calculate the urban form variable. Third, we consider the problems of sorting and endogeneity. Finally, we investigate other measures of urban form and examine the extensive margin of travel.

5.1. OLS estimations

Table 2 reports the results of OLS regressions of driving on urban form in US MSAs. Our unit of observation is a household described by the 2008 NHTS. In every column, our dependent variable is the log of household VKT, reported in the second row of Table 1. In all specifications, our measure of urban form is the log of 10-kilometer density, also as described by Table 1.

In column 1, we regress log annual household VKT on the log of density to find an elasticity of −8.7%. Households in locations with a 10% higher density drive 0.87% less and a one-standard deviation increase in density within 10 km is associated with a 0.11 standard deviation decrease in VKT. At the sample mean, this represents about 3,300 km annually. Because the estimated coefficient of density is stable across

Table 2
Driving and density, baseline OLS estimations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log 10-km density	−0.087[a]	−0.098[a]	−0.089[a]	−0.093[a]	−0.091[a]	−0.12[a]	−0.082[a]	−0.075[a]
	(0.0024)	(0.0020)	(0.0083)	(0.0089)	(0.013)	(0.0075)	(0.0051)	(0.0050)
White/Asian		0.019[b]			0.020[c]		0.024[b]	0.020[b]
		(0.0088)			(0.0100)		(0.0094)	(0.0090)
Share female		−0.26[a]			−0.26[a]		−0.26[a]	−0.26[a]
		(0.010)			(0.012)		(0.011)	(0.010)
log household size		0.49[a]			0.49[a]		0.49[a]	0.49[a]
		(0.011)			(0.011)		(0.011)	(0.012)
Single		−0.24[a]			−0.24[a]		−0.24[a]	−0.24[a]
		(0.012)			(0.013)		(0.013)	(0.013)
Age		0.045[a]			0.044[a]		0.044[a]	0.044[a]
		(0.0010)			(0.00099)		(0.00098)	(0.0011)
Age^2 (/1000)		−0.51[a]			−0.50[a]		−0.50[a]	−0.50[a]
		(0.0098)			(0.0095)		(0.0097)	(0.010)
log income		0.26[a]			0.25[a]		0.25[a]	0.25[a]
		(0.0043)			(0.0054)		(0.0053)	(0.0050)
Education		0.10[a]			0.095[a]		0.092[a]	0.091[a]
		(0.016)			(0.014)		(0.014)	(0.016)
$Education^2$		−0.014[a]			−0.012[a]		−0.012[a]	−0.011[a]
		(0.0024)			(0.0020)		(0.0020)	(0.0024)
log precipitation			−0.015		0.051		−0.13[b]	−0.060
			(0.053)		(0.042)		(0.057)	(0.079)
log precipitation sd			0.025		−0.036		0.11[c]	0.090
			(0.069)		(0.049)		(0.063)	(0.076)
Temperature			0.062[b]		0.049[b]		−0.080	−0.021
			(0.028)		(0.024)		(0.060)	(0.039)
Temperature sd			−0.043[b]		−0.033[c]		0.052	0.014
			(0.019)		(0.017)		(0.040)	(0.027)
Share higher educ.				−0.50[a]	−0.18		−0.25[a]	−0.30[a]
				(0.072)	(0.11)		(0.078)	(0.071)
Share higher $educ^2$				−0.020	−0.20[b]		−0.21[b]	−0.12
				(0.080)	(0.092)		(0.091)	(0.073)
log local income				0.68[a]	0.18[a]		0.23[a]	0.22[a]
				(0.014)	(0.033)		(0.014)	(0.015)
R^2	0.01	0.36	0.01	0.05	0.37	0.02	0.37	0.36
Observations	99,875	99,875	99,875	99,875	99,875	99,875	99,875	99,874

Notes: The dependent variables is log household VKT in all columns. All regressions include a constant including MSA fixed effects in columns 6 and 7 (275 MSAs) and county fixed effects in column 8 (837 counties). Robust standard errors in parentheses, clustered by MSA in columns 3–7, and by county in column 8. [a], [b], [c]: significant at 1%, 5%, 10%.

specifications, these magnitudes are relevant to most of the tables presented below. We also note that this elasticity is within the range of values estimated by past literature. Perhaps because we measure density more precisely, our elasticity is slightly larger in absolute value than the typical values of −4 to −5% estimated previously (e.g., Handy, 2005; Ewing and Cervero, 2010) as measurement error biases linear regression coefficients towards zero.

In column 2, we add household characteristics to our specification and estimate a slightly larger effect of density on household VKT, with an elasticity of −9.8%. White and Asian households drive about 2% more. Female households drive less. The coefficient of −0.26 implies that a single female is predicted to drive 23% less on average than a single male. Large households also drive more, but not proportionately so. The coefficients on log household size and the indicator for one-person households show that two-person households will drive about 30% more than one-person households. We also observe that VKT is concave in age. At age 20, an extra year of age is associated with 2% more driving. Then, VKT peaks around the age of 45 before declining. The elasticity of VKT with respect to income is large at around 26%. VKT increases with education (which is coded 1 to 5) for low levels of educational achievement and then decreases for the most educated households. Because the coefficients on households" characteristics are stable across specifications, we do not report or discuss them for subsequent tables.

In column 3, we consider geographic characteristics. Relative to column 1, the coefficient on density changes little. The results of this column indicate that VKT is higher where temperature is on average higher and varies less over the year. We find no significant effect of precipitation or its variation over the year. In other specifications like in column 7, we sometimes find that VKT is higher in places with less precipitation and more variation over the year.

In column 4, we consider neighborhood socio-economic characteristics. We find that driving declines with the share of university educated workers and increases with average local income. Because richer neighbourhoods are also on average denser, the coefficient on density also increases marginally in magnitude relative to the one estimated in column 1. In column 5, we consider all the controls together and estimate an elasticity of VKT with respect to density of −9.1%. Relative to column 4, we note that the magnitudes of neighbourhood characteristics drop sharply and lose significance. This is unsurprising. Richer and more educated households tend to live in richer and more educated neighbourhoods and the resulting co-linearity makes it difficult to separately estimate the effects of household and neighborhood income. We also note that despite the strong effects of neighbourhood and households characteristics, the coefficient on density barely changes. This is because these characteristics are only either modestly or weakly correlated with density and some cases, such as with education and income, their effects offset each other.

In column 6, we return to the specification of column 1 but also include a fixed effect for each Metropolitan Statistical Area (MSA). Estimating the elasticity of VKT with respect to density within MSAs yields a coefficient modestly larger in magnitude relative to column 1. This is because richer and more educated households, both drive more and tend to locate in denser MSAs. Consistent with this, including all the household, geographic, and neighbourhood characteristics in column 7 gives a coefficient on density close to that of column 5. This indicates that the between- and within-MSA variations in density are associated

Table 3
Robustness of baseline OLS estimations to measures of travel.

Dependent: Variable:	(1) stated km	(2) odometer km	(3) ind. day km	(4) dist. to work	(5) ind. day minutes	(6) speed	(7) number of trips	(8) mean trip distance
log 10-km density	−0.11[a]	−0.095[a]	−0.13[a]	−0.18[a]	−0.026[a]	−0.11[a]	0.014[a]	−0.15[a]
	(0.0054)	(0.0055)	(0.0066)	(0.0097)	(0.0036)	(0.0040)	(0.0020)	(0.0061)
R^2	0.42	0.43	0.18	0.11	0.12	0.14	0.33	0.10
Observations	93,602	71,742	83,313	86,387	85,996	82,849	83,313	83,313

Notes: All regressions include controls for household demographics, geography, local socio-economic characteristics, and MSA fixed effects. Robust standard errors clustered by MSAs in parentheses. [a], [b], [c]: significant at 1%, 5%, 10%. The dependent variables and explanatory variables of interest are in log in all columns except for the number of trips in column 7. Demographic controls include a white/Asian indicator, log income, log household size, a single indicator, age, age squared, gender, education, and education squared. Geographic controls include average precipitation and its standard deviation, and average temperature and its standard deviation. Local socio-economic controls include the share of residents with higher education and its square and log local income.

in a similar way with driving. We also note that despite the extra fixed-effects, the fit of the regression does not improve. We confirm below the lack of empirical content of these MSA fixed effects and provide more evidence that it is only the density surrounding households' residential location that determines their driving. The specification of column 7 is our benchmark OLS specification.[22] Finally, column 8 introduces a fixed effect for each of the 837 counties where metropolitan households are located. At −0.075, the coefficient on local density is marginally lower but statistically indistinguishable from our preferred coefficient in column 7 or from the coefficient obtained in column 1, the simplest estimation.

Our choice of explanatory variables in Table 2 controls for obvious determinants of household travel, like household demographics or the geography of where they live. We also include controls for the neighborhood socio-economic characteristics, in spite of the fact they are maybe correlated with urban form and capture some of its effect. Given our concern about the sorting of households on the basis of unobserved tastes for driving, we prefer the larger set of control variables.[23] As it turns out, once we control for basic household demographics, including further controls does not measurably affect the coefficient of urban form.

We postpone more detailed analysis, but we note that the elasticity of distance traveled with respect to density seems economically small. In Table 2, 10% increase in density corresponds to a less than 1% decline in distance traveled. Over the period 1990 to 2010, in only 1% of pixels housing an NHTS respondent does density increase by more than a factor of 4.8. Given the coefficient of −0.082 estimated in column 7 of Table 2, this implies that if the density of every residential location were to increase by this factor the corresponding decline in distance traveled is only of about 12%.

5.2. Robustness to measure of driving and urban form

In Table 3 we assess the stability of the results of Table 2 as we vary our dependent variable. In each column of this table we estimate a specification similar to that of column 7 of Table 2, with controls for households' demographics, neighbourhood socio-economic characteristics, and geography as well as a full set of MSA fixed effects. In column 1, we replace our preferred measure of VKT with a stated measure of VKT. We find a density elasticity of −11% instead of −8.2%. Measuring VKT through odometer readings by households in column 2, we estimate a density elasticity of −9.5%. Using a measure of daily VKT for individual drivers aggregated at the household level in column 3, the elasticity is again slightly larger at −13%. Using instead, distance to work in column 4 yields an even larger elasticity of −18%.

These elasticities for alternative measures of kilometers traveled are estimated on slightly different samples of households. In supplemental results we restrict attention to the about 37,000 households for whom we observe our preferred measure of travel and the four alternatives from columns 1–4 of Table 3, we find the following elasticities: −9.2% for our preferred measures of travel, −12% for stated miles, −11% for odometer miles, −14% for daily travel, and −18% for distance to work. The differences from the corresponding elasticities reported in Tables 2 and 3 are small.

We find some differences across different measures of travel, but note that these differences are small and that these measures are conceptually distinct. For instance, daily VKT is measured at the individual level whereas odometer VKT is measured for vehicles regardless of the number of household members who travel. Distance to work is more sensitive to local density. This is not surprising because commutes often take place when congestion is at its worst. Importantly, commutes represent 27% of household VKT and the density elasticity is −18% for commute distance. Hence, commutes account for $(0.27 \times 0.18)/0.092 \approx 53\%$ of the density elasticity of −9.2% that we estimate for all travel.

In column 5, our dependent variable is a measure of travel time, household daily travel minutes, that corresponds to kilometers traveled in column 3 and is directly measured by the survey. For this measure of travel time, we estimate an elasticity of −2.6%, much lower than for travel distance. In column 6, we use travel speed as the dependent variable and estimate an elasticity of −11%.[24] Although residents in denser locations travel fewer kilometers, their travel time is only marginally lower because travel is slower. In column 7, we use the number of trips as the dependent variable and estimate a small positive density elasticity of 1.4%. Finally, in column 8, we estimate an elasticity of mean trip distance to 10-kilometer density of −15%. This shows that the lower VKT of residents in denser locations is exclusively explained by shorter trips not by fewer trips. If anything, residents of denser locations tend to travel more often.

In Table 4, we assess the stability of the results of Table 2 as we vary our explanatory variable of interest. In column 1, we use 1-kilometer density to measure urban form instead of 10-kilometer density. Rela-

[22] In alternative specifications we also used the distance to the CBD as explanatory variable. Adding it in log to the specification of column 7 makes the coefficient on density marginally smaller in absolute value at −0.074. The elasticity of VKT with respect to distance to the CBD is small at 0.015.

[23] We experimented with many characteristics and included all those that are 'often' significant in the preliminary regressions we estimated. For instance, we include an indicator variables for households that are white or Asian. As can be seen in Table 2 below, this variable is often significant but the magnitude of its effects is small. We grouped white and Asian households because differences between them were minimal. Similarly we grouped all other minorities together because the differences between them were also minimal.

[24] Although our approach is very different from that developed in Couture et al. (2018), they estimate a comparable elasticity of travel speed with respect to population of −13% across the largest 100 US MSAs.

Table 4
Robustness of baseline OLS estimations to measure of density.

Sample restriction	(1) None	(2) None	(3) None	(4) None	(5) No NY	(6) No high density	(7) Non-MSA HH	(8) No high-VKT HH
Urban form:	1-km density	10-km pop. den.	10-km emp. den.	10-km land cover	10-km density	10-km density	10-km density	10-km density
	−0.067[a]	−0.083[a]	−0.065[a]	−0.055[a]	−0.080[a]	−0.078[a]	−0.081[a]	−0.067[a]
	(0.0036)	(0.0052)	(0.0046)	(0.0033)	(0.0045)	(0.0035)	(0.0046)	(0.0050)
R^2	0.37	0.37	0.37	0.36	0.37	0.37	0.38	0.34
Observations	99,875	99,875	99,870	99,423	94,970	74,864	26,328	90,662

Notes: All regressions include controls for household demographics, geography, local socio-economic characteristics, and MSA fixed effects. Robust standard errors clustered by MSAs in parentheses. [a], [b], [c]: significant at 1%, 5%, 10%. The dependent variables and explanatory variables of interest are in log in all columns. Demographic controls include a white/Asian indicator, log income, log household size, a single indicator, age, age squared, gender, education, and education squared. Geographic controls include average precipitation and its standard deviation, and average temperature and its standard deviation. Local socio-economic controls include the share of residents with higher education and its square and log local income.

tive to the −8.2% elasticity we estimate with 10-kilometer density, the estimate here is modestly lower at −6.7%. Columns 2 and 3 use the number of residents and the number of jobs within a 10 km radius to estimate comparable elasticities. We estimate a smaller elasticity in column 4 when using the share of developed land within a 10 km radius as measure of urban form. We return to these measures below when we consider several measures of urban form in the same regression.

In columns 5 to 8, we consider various sample restrictions to confirm that our results are not driven by a small subgroup of locations or drivers. In column 5, we estimate our preferred OLS estimation of column 7 of Table 2 without the New York MSA. Although travel behavior in New York is dramatically different from the rest of the country in many ways, surprisingly, the elasticity of VKT with respect to density is unchanged when we exclude it. In results not reported here, we estimate the same specification for only the New York MSA and obtain an elasticity of −14%. In column 6, we eliminate all observations in the top density quartile and still estimate an elasticity of −7.8%. In column 7, we consider only the non-MSA residents who are excluded from most of our specifications and estimate an elasticity of −8.1%. Finally, in column 8, we eliminate the 10% of households with the highest VKT. Collectively, these households are responsible for more than 20% of aggregate VKT. As these high VKT households are more often located in low density areas, we are bound to estimate a lower elasticity of VKT with respect to density. We do but, interestingly, the change is modest. We still estimate an elasticity of −6.7%.

Overall, these results suggest that our findings are broadly consistent across a variety of measures of driving and locations, but that particular measures of driving may be more or less sensitive to urban form.

5.3. Sorting

We now turn attention to the possibility that households and individuals in dense areas are different from those in less dense areas.

Prior investigations of urban form and driving considered the possibility of sorting into high-density residential areas in the context of Heckman selection models (see in particular Cao et al., 2009, for a full survey devoted to this type of estimation in our context). We report some results in Appendix B and note that this type of approach estimates coefficients for density that are only modestly larger in magnitude. However, a problem with this type of estimation is that it is not based on an exclusion restriction. That is, identification does not rely on a variable that would explain residential density but be otherwise uncorrelated with VKT. Instead, the possible sorting of households into high-density neighborhoods is identified entirely from assumptions about functional form. As a result, it is unclear how we should interpret the elasticities of VKT with respect to density estimated with this type of approach.

Our second approach to selection also follows standard ideas in the literature. In Table 2, we control for an increasingly rich set of observable individual characteristics. Intuitively, if such controls change the estimate of the coefficient of interest, then we worry that other unobserved variables might also be important. We see in Table 2 that this does not occur. Oster (2017) refines this intuition and points out that observed control variables do not generally inform us about the importance of unobserved controls unless the observed controls improve the R^2 of the regression. In addition, Oster (2017) provides a parametric test for bias caused by sorting on unobservables, conditional on an assumption about the extent to which unobserved controls are 'like' observed controls. In a nutshell, this test compares the coefficient of interest and the R^2 of a regression with added controls to one without. It rejects sorting when the difference between the two estimated coefficients of interest is small relative to the change in R^2. Performing this test on the regressions of columns 2 and 5 suggests that unobserved controls must behave very differently from observed controls in order to bias our estimates while columns 3 and 4 are uninformative about this issue.

Our third strategy for addressing the possibility of sorting revolves around variants of equation (15). Equation (15) can be implemented by regressing household vehicle-kilometers traveled on three explanatory variables of interest; the 1990 level of the density within 10 km, the change in this measure between 1990 and 2010, and the interaction of the change in density and a measure of mobility. Consistent with the discussion of Section 2, we proxy for the mobility rate in a given neighborhood with the mean tenure of a resident in the survey respondent's home cell.[25] We multiply by minus one so that increases in our proxy correspond to increases in mobility.

As discussed in Section 3, equation (15) offers two parametric tests of sorting. One of these tests involves the coefficient of the interaction of a mobility proxy with the change in density, and the second involves the difference between the coefficients of the level and of the change in density.

All of the specifications in Table 5 contain our three explanatory variables of interest. In addition, column 1 also includes the controls from our preferred specification in column 7 of Table 2 (household, neighbourhood and geographic characteristics, and MSA fixed effects). In order to address the possibility that driving behavior varies with tenure, column 2 also controls for the level of the mobility proxy. This specification closely approximates equation (15) and is our preferred specification. Column 3 also controls for the mobility rate interacted with the initial level of density. Column 4 restricts attention to bottom and top

[25] Our information on residential tenure comes from the ACS block group data (National Historical Geographic Information System, 2010). We impute this variable to grid cells as described in Section 4.

Table 5
Selection and mobility using information about local mobility measured through the tenure length of local residents.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Period	90 to 10	90 to 10	90 to 10	90 to 10	90 to 10	90 to 10	00 to 10	00 to 10
Household sample	All	All	All	Big Δ	Small Δ	Age <50	All	All
Initial log 10-km density	−0.080[a]	−0.075[a]	−0.046[a]	−0.084[a]	−0.069[a]	−0.080[a]	−0.076[a]	−0.076[a]
	(0.0052)	(0.0055)	(0.014)	(0.0073)	(0.0060)	(0.0072)	(0.0055)	(0.0054)
Δ log 10-km density	−0.12[a]	−0.063[b]	−0.030	−0.061	−0.056	−0.033	−0.014	−0.014
	(0.025)	(0.029)	(0.033)	(0.054)	(0.060)	(0.035)	(0.043)	(0.044)
Mobility ×Δ log density	0.0077[b]	−0.0033	−0.00059	0.0014	−0.0048	0.0023	−0.00014	−0.00014
	(0.0034)	(0.0036)	(0.0038)	(0.0067)	(0.0072)	(0.0044)	(0.0057)	(0.0057)
Mobility rate		−0.0099[a]	−0.040[a]	−0.0056	−0.016[a]	−0.010[a]	−0.010[a]	−0.010[a]
		(0.0029)	(0.013)	(0.0043)	(0.0044)	(0.0033)	(0.0027)	(0.0027)
Mobility × log density			0.0026[b]					
			(0.0011)					
Past Δ log 10-km density								−0.0017
								(0.022)
F-test 1 p-value	0.061	0.0020	0.24	0.82	0.44	0.13	0.0011	0.0055
F-test 2 p-value	0.073	0.65	0.54	0.71	0.88	0.16	0.14	0.15
R^2	0.37	0.37	0.37	0.36	0.37	0.26	0.37	0.37
Observations	99,875	99,875	99,875	46,942	48,939	39,253	99,875	99,875
Number of MSA	275	275	275	263	272	275	275	275

Notes: The dependent variable is log household VKT in all columns. Mobility is measured as - average tenure length in of residents of the same home cell (sample mean, 10.3 years and standard deviation 2.4 years). All regressions are estimated with OLS and include MSA fixed effects with demographic controls (a white/Asian indicator, log income, log household size, a single indicator, age, age squared, gender, education, and education squared), geographic controls (average precipitation and its standard deviation, and average temperature and its standard deviation) and local socio-economic controls (the share of residents with higher education and its square and log local income). Robust standard errors clustered by MSA in parentheses. [a], [b], [c]: significant at 1%, 5%, 10%. F-test 1 is a joint test of the equality of the coefficients on initial log 10-km density and Δ log 10-km density and of the coefficient on mobility ×Δ log density being zero. F-test 2 is a test of the equality of the coefficients on Initial log 10-km density and Δ log 10-km density.

quartile of density growth in a 10-kilometer radius (excluding the top and bottom percentiles). Column 5 considers the complementary sample of households located in locations at the second and third quartile of density change. Column 6 restricts attention to survey respondents with household age below 50. Columns 7 and 8 consider the ten-year periods from 1990 to 2000 and from 2000 to 2010.

In every case, we find that the coefficient on initial density and that on the change in density are statistically close. Except for columns 1 and 8, the coefficient on the change in density is less than a standard deviation from the coefficient on density. With this said, point estimates are different and, from equation (15), the difference between these two coefficients is α_1, our measure of selection. Hence, while we cannot, in a majority of cases, reject the hypothesis that $\alpha_1 = 0$, point estimates suggest it is negative. In our preferred specification in column 2, we have $\alpha_1 = -0.075 - (-0.063) = -0.012$ and $\beta = -0.063$, so that sorting accounts for about one sixth of the effect of density on driving. In equation (15), we can also implement the second test for $\alpha_1 = 0$ by rejecting the hypothesis that the coefficient of Mobility $\times \Delta$ log(10-kilometer density)= 0. In every specification, we see that this coefficient is small, precisely estimated and usually indistinguishable from zero. This also suggests that α_1 is small.

We note that, in the same spirit as equation (15), we can also compare the coefficient on initial density in high-mobility locations (column 4) and low-mobility locations (column 5). In addition, we can even compare the coefficient obtained when estimating our preferred specification on a sample of more mobile residents (those below 50 as in column 6) to the overall sample in column 2. In both cases, the differences are close to zero and the coefficients are precisely estimated.

More generally, and in light of the hypothesis tests developed in Section 2, Table 5 suggests that as density changes the driving behavior of people who leave is not statistically distinguishable from that of the people who arrive. With this said, point estimates indicate a modest amount of sorting.

In the remainder of this section, we report a number of robustness tests for this result. First, appendix Table 13 replicates our preferred estimation from column 2 of Table 5 under various sample restrictions, with a purely residential population based measure of density, and using alternative dependent variables. These results are consistent with our findings so far. Excluding high-density locations or high-VKT households makes no difference. Focusing more narrowly on more mobile households in locations facing greater changes in population or on less mobile households in more stable locations yields elasticities of VKT with respect to density that are of the same magnitude. Using only population instead of population and employment to measure density makes no difference. We also confirm the results of Table 3. That is, the elasticity of daily VKT is slightly larger than the annual measure, the elasticity of travel time is close to zero, and this difference is still explained by the difference in travel speed.

Second, appendix Table 14 presents a series of regressions that are identical to those of Table 5, except that we proxy for the mobility rate with the share of renters in the cell of the survey residents. These results are qualitatively similar to those of Table 5 except that the interaction terms are marginally larger and are estimated somewhat less precisely. In spite of this, these results suggest the same conclusion as does Table 5. That is, as urban form changes, the driving behavior of arrivals is like that of those who leave.

We next use age as a proxy for mobility. However, given that the relationship between age and residential mobility is unlikely to be linear, we use a vector of decadal age dummies to describe the age of drivers. Then, consistent with the intuition developed in equation (15), we interact these indicators with changes in urban form. We include these interactions in regressions that also contain the 1990 level of urban form and changes in urban form. Table 6 reports these results. Column 1 includes only the log level and change of density within 10 km of a survey respondent's home cell, along with an extensive set of control variables. Column 2 includes the interaction terms. Columns 3–8 repeat column 1 on a variety of subsamples. The results of this table are striking. In every specification the coefficient of the level and change in urban form are statistically indistinguishable and coefficients do not vary across specifications. This does not allow us to reject $\alpha_1 = 0$ in equation (15), and as above, in most specifications point estimates suggest that

Table 6
Sorting on age OLS estimations.

Household sample	(1) All	(2) All	(3) Age<50	(4) Age>60	(5) Big Δ	(6) Small Δ	(7) Big Δ Age<50	(8) Small Δ Age>60
log 10-km density 1990	−0.082[a]	−0.085[a]	−0.086[a]	−0.074[a]	−0.087[a]	−0.080[a]	−0.087[a]	−0.073[a]
	(0.0053)	(0.0063)	(0.0073)	(0.0053)	(0.0068)	(0.0068)	(0.0086)	(0.0076)
$\Delta_{1990-2010}$ log 10-km density	−0.071[a]	−0.068[a]	−0.080[a]	−0.058[b]	−0.091[a]	−0.093	−0.092[a]	−0.13
	(0.013)	(0.019)	(0.022)	(0.025)	(0.019)	(0.057)	(0.027)	(0.096)
Controls:								
Demographics	Y	Y	Y	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y	Y	Y	Y
Local socio-econ.	Y	Y	Y	Y	Y	Y	Y	Y
Decade indicators	N	Y	N	N	N	N	N	N
Decade × log density	N	Y	N	N	N	N	N	N
Decade × Δ log density	N	Y	N	N	N	N	N	N
F-test 1 p-value	.	0.0028	.	.	.	.	.	.
F-test 2 p-value	0.31	0.32	0.71	0.51	0.80	0.81	0.81	0.52
R^2	0.37	0.37	0.26	0.26	0.36	0.37	0.25	0.27
Observations	99,875	99,875	39,253	40,421	46,942	48,939	18,710	19,980
Number of MSA	275	275	274	274	263	272	247	257

Notes: All regressions include MSA fixed effects. Robust standard errors clustered by MSA in parentheses. [a], [b], [c]: significant at 1%, 5%, 10%. The dependent variables and explanatory variables of interest are in log in all columns. Demographic controls include a white/Asian indicator, log income, log household size, a single indicator, age, age squared, gender, education, and education squared. Geographic controls include average precipitation and its standard deviation, and average temperature and its standard deviation. Local socio-economic controls include the share of residents with higher education and its square and log local income. When decade effects are introduced, households in their 40s are used as reference. See Table 7 for the detailed results of column 2. F-test 1 is a joint test of the equality of the coefficients on log 10-km density in 1990 and Δ log 10-km density and of the coefficients on decade indicators interacted with Δ log density all being zero. F-test 2 is a test of the equality of the coefficients on Initial log 10-km density and Δ log 10-km density.

Table 7
Detailed results for column 2 of Table 6: age decade indicators.

Age	20–29	30–39	40–49 (ref.)	50–59	60–69	>70
Decade indicators	−0.057	0.087	0	0.034	−0.22[a]	−0.13
	(0.098)	(0.10)		(0.075)	(0.082)	(0.087)
Decade × log 10-km density 1990	0.0033	−0.0073	0	−0.0039	0.014[b]	0.0035
	(0.0079)	(0.0085)		(0.0060)	(0.0068)	(0.0067)
Decade × Δ_{90-10} log 10-km density	−0.010	−0.0097	0	0.027	0.022	−0.053
	(0.022)	(0.020)		(0.020)	(0.028)	(0.036)

Notes: This table reports the coefficients on decades of age, interactions between decades of age and log 10-km density in 1990, and interactions between decades of age and log density changes between 1990 and 2010.

α_1 is a small negative number. Table 15 in appendix provides further robustness checks for these results.

Table 7 reports the interaction terms for column 2 of Table 6. On the basis of equation (15) the coefficients of the last set of interaction terms, the interaction of decade of life with change in urban form, should inform us about α_1 for that subgroup. We see that these coefficients are all small relative to the effect of density on driving and are indistinguishable from zero. We note that the table includes a complete set of interaction as controls. We are concerned that driving behavior may vary by age or that relationship between driving and density was different in places with different initial demographics.

We have now completed five distinct tests of the role of sorting. First, we report the result of Heckman-type corrections for residential selection into high density. These estimates suggest that the relationship between urban form and driving does not reflect sorting of individuals across places on the basis of their propensity to drive. Second, in our OLS results, we control for observable characteristics. We find these controls have only a tiny effect on our estimates of the effect of density on driving and the more formal test of Oster (2017) indicates that unobservables are unlikely to bias our estimates. Finally, we also develop a parametric test for the role of sorting and implement it using three different proxies for the mobility rate of residents. In each case, we find little support for the idea that sorting is an important determinant of the relationship between density and driving.

5.4. Endogeneity

Table 8 reports the results of a series of instrumental variables estimations. These regressions are all variants of equation (16) in which we rely on permutations of three types of instruments. These instruments measure the share of the 10-kilometer disk surrounding a respondent's home cell that overlays an aquifer that can provide residential water. This variable is well known to predict urban form (Burchfield et al., 2006).[26] In addition, we construct variables measuring a respondent's exposure to earthquakes and landslides. These variables have a remarkably strong ability to predict surface employment and residential density, and it is not easy to see how they might influence driving through

[26] Note that our use of aquifers slightly differs from that in (Burchfield et al., 2006). In their work, widely available underground water in an entire metropolitan area is shown to cause low density and scattered development. In our analysis, we work at a finer geographical scale and compare areas with underground water, which enjoy more development, and areas without water, which are less attractive for development.

Table 8
IV regressions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log 10-km density	−0.13[b]	−0.100[c]	−0.12[a]	−0.076[a]	−0.080[a]	−0.075[c]	−0.069[a]	−0.075[a]
	(0.060)	(0.054)	(0.032)	(0.026)	(0.025)	(0.041)	(0.025)	(0.024)
Controls	N	Y	Y	Y	Y	Y	Y	Y
MSA effects	N	Y	Y	Y	Y	Y	Y	Y
Instruments:								
Aquifers	Y	Y	N	N	Y	Y	N	Y
Earthquakes	N	N	18	N	N	3	3	3
Landslides	N	N	N	Y	Y	N	Y	Y
Over identification p-value	.	.	0.14	0.27	0.39	0.60	0.38	0.45
First-stage statistic	202	32.6	24.2	81.3	82.4	29.5	87.6	83.8
Observations	99,874	99,874	99,874	99,874	99,874	99,874	99,874	99,874
Number of MSA	275	275	275	275	275	275	275	275

Notes: All regressions TSLS regressions with a constant. Controls are demographic controls (a white/Asian indicator, log income, log household size, a single indicator, age, age squared, gender, education, and education squared), geographic controls (average precipitation and its standard deviation, and average temperature and its standard deviation) and local socio-economic controls (the share of residents with higher education and its square and log local income). In column 3, we use all 18 values of earthquake intensity as dummy variables. In columns 6 to 8, we group them into three groups (intensity below 2, between 3 and 14, and above 15.) [a], [b], [c]: significant at 1%, 5%, 10%. Robust standard errors clustered by county in parentheses. Clustering is by county to have a sufficient number of clusters to compute robust covariance matrices more reliably than when clustering by MSA. The dependent variables and explanatory variables of interest are in log in all columns. We do not report first-stage results here given that we use 25 different variants for our instruments (most of them to measure earthquakes). We nonetheless note that lower exposures to landslide or earthquakes and higher presence of aquifers are (conditionally) positively associated with greater density within 10 km.

Table 9
Driving and urban form, extended OLS estimations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log 10-km density	−0.082[a]	−0.082[a]				−0.088[a]	−0.089[a]	−0.084[a]
	(0.0055)	(0.0065)				(0.0061)	(0.0054)	(0.0059)
log 10-km job ratio	−0.0020							
	(0.0066)							
log 10-km corrected density		0.0036						
		(0.0065)						
log 10-km land cover			−0.0051					
			(0.0064)					
log 10-km population			−0.050[a]					
			(0.013)					
log 10-km employment			−0.024[a]					
			(0.0068)					
log 10-km weighted density				−0.043[a]				
				(0.0024)				
log 1-km density					−0.026[a]			
					(0.0066)			
log 1-to-5 km density					−0.044[a]			
					(0.0099)			
log 5-to-10 km density					−0.0042			
					(0.0078)			
log 10-to-25 km density					−0.0091			
					(0.0080)			
log 1-km roads						−0.00086[c]		
						(0.00048)		
log 25-km roads						0.018[b]		
						(0.0082)		
log 25-km arterials							0.022[a]	
							(0.0057)	
log 25-km highways								0.0015
								(0.0011)
R^2	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Observations	99,870	99,423	99,423	99,861	99,861	99,875	99,875	99,875
Number of MSA	275	275	275	275	275	275	275	275

Notes: The dependent variable is log household VKT in all columns. All regressions are OLS regressions with MSA fixed effects. Controls are demographic controls (a white/Asian indicator, log income, log household size, a single indicator, age, age squared, gender, education, and education squared), geographic controls (average precipitation and its standard deviation, and average temperature and its standard deviation) and local socio-economic controls (the share of residents with higher education and its square and log local income). Corrected density in column 2 measures residential population and employment within a 10-kilometer radius relative to developed land. Weighted density in column 4 is a weighted sum of density within one kilometer (weight=1), density from one to five kilometers (weight=0.5), density from five to 10 km (weight =0.25), and density from 10 to 25 km (weight=0.125). Robust standard errors clustered by MSA in parentheses. [a], [b], [c]: significant at 1%, 5%, 10%.

any other channel given that we control extensively for local geographic and socio-economic characteristics.[27]

In column 1 we present an instrumental variables regression using our aquifers instrument but do not include other controls. In the second column, we add MSA indicators and the same long list of controls that we use in column 7 of Table 2. In the subsequent columns we experiment with the different instruments and with permutations of these instruments. The coefficient of density is stable across specifications. In every case our instruments are not weak according to conventional tests, and in regressions including more than one instrument, we comfortably pass over-id tests.

Most importantly, coefficient estimates are statistically indistinguishable from those in our table of OLS estimations. This suggests that omitted variables correlated with driving and urban form are not causing economically important bias in our estimates of the relationship between urban form and driving.[28]

5.5. Other urban form variables

On the basis of our work so far, it appears that neither sorting nor omitted variables cause bias in OLS estimates. Given this, we now turn to an investigation of the effects of different measures and spatial scales of urban form on driving using OLS regressions.

Table 9 reports results for a series of 'horse races' between measures of urban form. Three main conclusions emerge from this table. The first is that, although population and employment appear to play a role in explaining VKT, once we use our preferred measure of density of both jobs and residents within 10 km, other measures such as the ratio of jobs to residents have no measurable effect on VKT despite small standard errors. This conclusion holds more broadly than for the specifications we reported here. We experimented extensively with measures of job vs. residential locations. The effect we estimate for our preferred measure of 10-kilometer density is robust to the inclusion of many alternative measures of urban form and none of these alternative measures of urban form appears to systematically affect VKT.

Our second conclusion concerns roads. We estimate a small positive association between roads within a 25-kilometer radius of a household's place of residence and household VKT. We acknowledge that roads may be simultaneously determined with VKT. This said, we note that the small effects of roads that we estimate are conditional on density and many other variables. Such small effects are not inconsistent with new major arterials and highways eliciting a lot of traffic as households may choose to locate closer to roads (Baum-Snow, 2007) or substitute across roads.[29]

Our third conclusion is that most of the effect of density takes place within 10 km of a household's place of residence. The results of column 5 are even suggestive that it is density within 5 km that is most important. In spite of this, when used 'alone' 10-kilometer density is often more precisely estimated than 5-kilometer density, and so we rely more heavily on 10-kilometer density in reported results.

In regressions not reported here, we have also used the MSA fixed effects estimated in our preferred OLS regression from column 7 of Table 2 and regressed them on variables that describe MSAs. We found no effect of MSA population, area, education, income, or geography. We also found no effect of measures of MSA employment concentration, residential concentration, and mismatch between jobs and residents. We found weak effects for some measures of segregation and the share of manufacturing employment. As we experimented with a large number of MSA characteristics, we expect the coefficients of a small proportion of them to be significant. We interpret this large majority of insignificant coefficients as an absence of MSA effects after controlling for the characteristics of households and their immediate surroundings. This absence of metropolitan effect is consistent with the fact that the insertion of MSA fixed effects in Table 2 does not improve the R^2 of the regressions. That most of the effect of urban form on VKT should take place within a reasonably short range may not be surprising given that mean trip distance is slightly less than 13 km in our data.[30] A full exploration of this issue is beyond the scope of this paper as it would involve revisiting the data and specifications of previous contributions using our measures of density.

5.6. Non-linearities and mode choice

We now return to a feature first apparent in Fig. 3, the possible non-linearity of the relationship between log household VKT and log density. Column 1 of Table 10 enriches our baseline OLS regression of column 7 of Table 2 with a quadratic term for log density and suggests that the effects of density within a 10 km radius becomes stronger at higher levels of density. Economically, the increase in the magnitude of the density elasticity of VKT is modest. The coefficient on squared log density of −0.070 implies a 2.3 percentage point difference between the bottom and top density deciles relative to a baseline estimate of −8.2%. This finding is confirmed in column 2 where we instead consider two density thresholds. The results of this column indicate that the density elasticity of VKT is slightly less than one percentage point higher for densities above the 95th percentile and another 1.5 percentage point higher for densities in the top percentile (where only 1% of households in our sample reside). Although there is a 'high density' effect, it remains modest.

In the rest of Table 10, we turn to the extensive margin of urban travel and examine the possible substitution across modes. For this, we return to the information about individual trips and consider three types of trips, privately-owned vehicles, any form of transit, and walking or biking trips. The results from our logit estimations indicate that the propensity to use privately-owned vehicles for a trip declines with density but most of the effect appears concentrated at the top density percentile. For transit, the relationship is non-monotonic and this mode of transportation is more prevalent at low and high density. Finally, the share of walking and biking trips increases with density but the relationship is only significant in the top five density deciles.

These results about the extensive margin of travel do not alter our conclusions so far. For residents of US metropolitan areas, the share of trips by privately-owned vehicles is about 89% while the share of transit is less than 2%. Biking or walking trips represent about 9% of trips (but only a trivial share of kilometers travelled). Although the coefficient (an odds ratio) on privately-owned vehicles at the top centile of

[27] One may imagine that these variables may affect VKT indirectly through the road infrastructure. In results not reported here, we verify that adding measures of the lane kilometers of major roads does not affect our results. This is consistent with the weak effect of nearby highways and major roads on household VKT uncovered below. We also verify that our results are robust to the distance to the central business district.

[28] Including measures of topography in our results does not change the coefficients on urban form variables in either OLS or IV results. However, it does change our first stage. In particular, our underground geology variables do not generally pass weak instrument tests if we include topographical variables as controls.

[29] In Duranton and Turner (2011), we regress log highway VKT on highway lane kilometers and estimate an elasticity close to unity. Despite their apparent similarity with the estimations reported here, the regressions of Duranton and Turner (2011) are very different because they consider VKT for road segments, not households. Three features are associated with this difference. First, highway VKT represents only about a quarter of aggregate VKT. Second, there is likely a lot of potential substitution between local roads and highways. Third, we only consider driving by people who live nearby, thus ignoring VKT by households who live further away, households who relocate, and commercial traffic.

[30] Unlike us, much of extant literature finds that sizeable effects of other measures of urban form on driving (e.g., see Stevens, 2017, for a summary of these findings). The reasons for this important difference are unclear. We suspect they might be due to our use of more stringent controls for household and local characteristics.

Table 10
Non-linearities and mode choice.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regression	OLS	OLS	logit	logit	logit	logit	logit	logit
Dep. var.	VKT	VKT	trip POV	trip POV	trip transit	trip transit	trip walk/bike	trip walk/bike
log 10-km density	−0.0056	−0.078[a]	−0.037[a]	0.0024	−0.041[a]	−0.089[a]	0.046[a]	0.014[a]
	(0.030)	(0.0041)	(0.0042)	(0.0043)	(0.010)	(0.010)	(0.0047)	(0.0048)
log 10-km density^2	−0.0070[b]							
	(0.0029)							
log 10-km density, above 95th pctl		−0.0083[a]		−0.037[a]		0.030[a]		0.031[a]
		(0.0028)		(0.0021)		(0.0063)		(0.0022)
log 10-km density, above 99th pctl		−0.015[a]		−0.081[a]		0.13[a]		0.057[a]
		(0.0048)		(0.0032)		(0.0076)		(0.0035)
R^2	0.37	0.37	0.03	0.03	0.12	0.13	0.03	0.03
Observations	99,875	99,875	837,647	837,647	827,685	827,685	837,606	837,606
Number of MSA	275	275	275	275	211	211	266	266

Notes: The dependent variable is log household VKT in columns 1 and 2, a trip indicator variable taking a value of 1 for trips with privately owned vehicles in column 3 and 4, a trip indicator variable taking a value of 1 for transit trips in columns 5 and 6, and a trip indicator variable taking a value of 1 for walking or biking trip in columns 7 and 8. OLS regressions in columns 1 and 2 and logit regressions in columns 3–8. Odds ratios reported for all logit regressions. All regressions include MSA fixed effects. Controls are demographic controls (a white/Asian indicator, log income, log household size, a single indicator, age, age squared, gender, education, and education squared), geographic controls (average precipitation and its standard deviation, and average temperature and its standard deviation) and local socio-economic controls (the share of residents with higher education and its square and log local income). Standard errors in parentheses. [a], [b], [c]: significant at 1%, 5%, 10%.

residential density in column 4 of Table 10 and that on transit in column in column 6 should not be dismissed as tiny, it is important to keep in mind that even for the 1% densest households, the share of transit trips is only 6.4%. At best, the mode switches we observe at high density can only explain the higher elasticity of VKT with respect to density that we observe at the same high levels of density in columns 1 and 2 of Table 10.[31]

6. Discussion

6.1. Using the model: Driving and welfare

We now return to our model and use it to perform a series of simple calculations that allow us to highlight the consistency of our findings with a range of previous results and provide some suggestive conclusions regarding their welfare implications.

First consider the case when $\alpha_1 = 0$ and there is no sorting. We also find no evidence of an important role for unobserved local characteristics. This suggests that δ, the unobserved propensity to drive, is uncorrelated with X, conditional on controls. Together these two conditions imply that the coefficient β estimated from a regression of log VKT on log density identifies $-\frac{\phi-\zeta\rho}{1-\rho+\phi}$ as per equation (7) of our model. The OLS estimate of β in column 7 of Table 2 is −0.082. Taking alternative measures of travel distance, the first 3 columns of Table 3 estimate slightly larger magnitudes for β between −0.095 and −0.13. To ease calculation, say $\beta = -0.1$, and we have,

$$\beta = -\frac{\phi - \zeta\rho}{1 - \rho + \phi} = -0.1 \,. \tag{18}$$

By dividing travel distance in equation (7) by mean trip distance in equation (5), we obtain the number of trips. Hence, regressing the log number of trips on log density provides an estimate of $\zeta + \beta$, where ζ is the accessibility elasticity. Column 7 of Table 3 provides such an estimate. It suggests that $\zeta = -\beta + 0.014 = 0.114$.[32]

From equations (6) and (7), regressing speed – an inverse measure of travel cost – on density, provides an estimate of $(1 + \beta)\phi$. Hence, our value of −0.1 for β and the estimated density elasticity of speed of −0.107 in column 6 of Table 3 imply $\phi = 0.119$ for the congestion elasticity. Knowing β, ζ, and ϕ, it is now easy from equation (18) to provide a value for ρ, the concavity of utility: $\rho = 0.5$.

We note that the implied values of the accessibility and congestion externalities, ζ, and ϕ, are not sensitive to the exact choice of β. By contrast, the implied value of ρ is sensitive to β. A value of −0.09 for β implies $\rho > 1$ whereas a value of −0.11 implies $\rho < 0$. This is because the value of ρ in equation (18) results from dividing a small numerator by a small denominator.

While the absence of sorting is a good first-order approximation, we now verify that considering sorting explicitly does not affect these conclusions. Hence, we now consider situations with sorting. We assume above that $\theta = \overline{\theta}\nu$. Using this in equation (7) and the parameterization of sorting in equation (12) implies that the density elasticity of VKT is now $\frac{\alpha_1}{1-\rho+\phi} + \beta = -0.1$. We can obtain an estimate of the sorting term $\frac{\alpha_1}{1-\rho+\phi}$ from the difference between the coefficient on density and that on the change in density in Table 5. Our preferred estimate from column 2 of Table 5 indicates $\frac{\alpha_1}{1-\rho+\phi} = -0.012$. Hence we have $\beta = -0.112$ when we consider sorting instead of −0.1 when we do not. From equations (4) and (7), the density elasticity of the number of trips is now $\frac{\alpha_1}{1-\rho+\phi} + \beta + \zeta$. This implies $\zeta = 0.126$. From equations (5) and (7), the density elasticity of speed is now $\left(1 + \beta + \frac{\alpha_1}{1-\rho+\phi}\right)\phi$ which leaves the value of $\phi = 0.119$ unchanged as sorting affects travel distance and travel time in the same manner and thus disappears when estimating speed as function of density.[33]

[31] Interestingly, in results not reported here where we additionally control for trip distance in the same logit estimations, most of the effect of density disappears as walking and biking trips are in their overwhelming majority short trips while transit trips are either short or long. This is consistent with the notion that density reduces driving by increasing accessibility as documented in Table 3.

[32] We can also obtain a value $\zeta + \beta$ by taking the difference between the density elasticity of mean trip distance in column 8 of Table 3 and the density elasticity of daily VKT in column 3 of the same table. These two measures are directly comparable as they rely on the same measure of VKT. They imply that the sum $\zeta + \beta$ is very much the same: $0.151 - 0.134 = 0.017$ instead of 0.014 when this quantity is estimated directly in column 7 of table 3.

[33] We note that ρ is no longer separately identified from α_1 in this context unless further assumptions are being made.

While there is a large literature that estimates congestion effects through traffic flows and traffic speed (Small and Verhoef, 2007), most of it is concerned with estimating effect of the (endogenous) number of vehicles on traffic speed for a particular segment of roads or groups of road segments. Attempts to measure congestion for an area depending on its population are extremely rare. Couture et al. (2018) estimate the effect of MSA vehicle travel time on a measure of MSA speed and find an elasticity of −0.13 for the largest 100 US MSAs. With the caveat that MSA population and the density of residents and workers are different objects, we nonetheless note that this estimated value of ϕ of 0.13 in Couture et al. (2018) is very similar to our implied value of 0.119 despite a very different methodology.

We know of no alternative estimate of the accessibility elasticity ζ in the literature that could be directly compared with ours. Couture (2014) estimates the (constant) elasticity of substitution between restaurants using a logit model of travel demand. His framework imposes a constant trip time, which is consistent with the extremely small elasticity we estimate in Table 3. His estimates of the elasticity of substitution are about nine which are consistent with accessibility benefits associated with the number of restaurants of about (9/8-1=)0.125. Although this comparison is somewhat of a stretch, this value is remarkably close to our estimate of $\zeta = 0.126$ with sorting.

Although they do not explicitly model travel behaviour, Ahlfeldt et al. (2015) estimates the consumption benefits from greater population density with a structural model that they implement using detailed data for the city of Berlin. Their structural model estimates an elasticity of block-level amenities with respect to a discounted measure of nearby residential density of 14%. This measure of consumption spillover is probably best interpreted as a measure of the importance of accessibility to nearby amenities and goods. To repeat, this does not directly correspond to our measure of accessibility ζ but it is nonetheless suggestive of a similar magnitude.

As another check on the consistency of our results, we can return to equilibrium utility as given by equation (8). As made clear by this equation, the elasticity of utility (which maps directly into consumption expenditure) with respect to density is $\frac{\rho}{1-\rho+\phi}[\zeta(1+\phi)-\phi] \equiv b$. Using our implicit value of the congestion elasticity ϕ of 0.119, our implicit value of the accessibility elasticity ζ of 0.126, and our preferred value of $\rho = 0.5$, we obtain an elasticity of utility with respect to density below 0.02. This implies that equilibrium utility is fairly insensitive to density.[34] In turn, this is consistent with little sorting being detected in our data.

The simple model proposed in Section 2 was first used to derive an empirical specification and discuss identification concerns. In this section, we connect our empirical results to our model. This leads to the following conclusions. Our empirical results imply estimates of the accessibility and congestion elasticities which are consistent with previous literature. In line with our empirical results about the weakness of sorting, these structural parameters also imply weak effects of density on utility as the accessibility and congestion externality essentially offset each other.

6.2. Some simple equilibrium implications

Policies that aim to both increase population and employment density and reduce driving are often referred to as 'smart growth' policies. These policies are hard to evaluate directly as they do not generally set explicit density targets and work indirectly through a wide range of instruments. We note nonetheless they generally emphasise high-density new developments and growth containment through the preservation of less developed land.[35] We also note that these policies only modestly affect urban form. Despite its widely recognized (and often praised) adoption of smart growth policies, the metropolitan area of Portland saw its population density weighted at the census tract level go up by only 9% between 2000 and 2010. Portland remains the 24th densest metropolitan area in the US. For comparison, the metropolitan areas of San Francisco and New York are about three and seven times denser, respectively. To make our analysis more transparent, we consider densification policies that are simpler and more drastic.

Table 11
Driving and population by density decile.

(1) Decile	(2) Area share (%)	(3) Density	(4) VKT pp	(5) Area VKT share (%)	(6) Area VKT (10^9 km)
Continental US:					
1	83.26%	6.25	19,329	12.2%	620.86
10	0.21%	5,421.04	12,497	7.9%	401.44
MSA only:					
5	1.94%	816.19	15,175	9.9%	391.14
9	0.82%	2,741.09	13,779	9.0%	355.14

Notes: Top panel describes first and tenth density deciles of US population. 2010 census and 2008 NHTS populations are 3.08 million and 321,000, the total area of the continental US is 7.03 million km^2, and total NHTS VKT is 5.08 trillion kilometers. The bottom panel describes the fifth and ninth density deciles of MSA population. Census and NHTS MSA populations are 2.47 million and 257,000, the total area of the continental US MSAs is 1.66m km^2, and total NHTS VKT is 3.94 trillion kilometers.

Table 11 describes the way that people, driving and density are distributed. The top panel describes the continental US and the bottom panel restricts attention to the approximately 20% of area and 80% of population within MSA boundaries, the sample on which our regressions are primarily based. The rows of each panel describe 'density deciles of population'. For example, the first row of the top panel describes the 10% of the NHTS population living in the least dense parts of the US while the second row describes the 10% of the NHTS population living in the densest parts of the country. These are the subsets of the population that live in the white and black regions of Fig. 2a.[36]

Calculating density deciles of population requires calculating threshold values of density that divide the NHTS population into tenths. The second column of the table describes the share of land area occupied at the densities intermediate between these thresholds. For example, 83% of US land area is occupied at lower densities than the threshold density for the bottom density decile of the NHTS population. Moving across columns to the right; the average density of these pixels is 6.25 people or jobs per square kilometer, the average travel per person for NHTS people living in these pixels is 19,329 vehicle-kilometers, and the population of this decile accounts for 12% of all driving in the NHTS, as measured by household odometer readings. Finally, column 7 gives aggregate driving in each decile in billions of kilometers per year.

Table 11 permits calculations to assess the impact of policies to change density on aggregate driving. For example, consider a policy which relocates the bottom density decile of US population into an area whose density is equal to the average density occupied by the top density decile of population. To implement such a policy we would take population dispersed across 83% of the country's land area and settle them in about 0.2% of the country's land area, concentrating the population resident in the white area of the map in Fig. 2a into an area the size of the barely visible black area. From column 3 of Table 11, this involves an 867 fold increase in density, from 6.25 to 5,421. From our estimate in column 7 of Table 2, this results in about a $(1-(5421/6.25)^{0.082}=)$43%

[34] As a caveat, we note that the elasticity of utility with respect to density increases with ρ. For $\rho = 0.8$, it is equal to 0.06 and rises to 0.18 for $\rho = 1$.

[35] https://www.smartgrowthamerica.org/ , June 12, 2017.

[36] While most of our results so far were derived for MSA households to be able to include MSA fixed effects, we verified that similar results are obtained for non-MSA households including in column 7 of Table 4. We now work with the entire country for our counterfactual computations.

decrease in driving for this decile of the population. Since this decile of population accounts for about 12% of total driving, this gives about a 5.1% decrease in aggregate driving.

This policy is particularly drastic as it involves increasing density though a massive reduction in land area. More plausible densification policies arguably involve reallocating a part of the population from less dense areas to denser areas. We consider for instance a policy that moves 1% of the MSA population and employment from the area inhabited by the fifth population decile of density to the ninth. Although only 1% of the population moves, the entire 20% of the MSA population in the source and destination regions experiences a change in density. Thus, to calculate the aggregate change in VKT we must calculate the change in aggregate driving for three groups; the 9% of the population that stays in region five, the 10% of population initially in region nine and the 1% of the population that moves from region five to region nine. The group that stays in region five is initially responsible for 0.9×391.14 billion VKT. They experience a 10% reduction in density. Using our preferred density elasticity of −0.082, this causes driving to increase by a factor of 1.0087, an increase of 3.06 billion VKT. The population of region nine initially drives 355.14 billion VKT. They experience a 10% increase in density which causes their driving to decrease by a factor of 0.9922, for a total decrease of 2.77 billion VKT. Finally, the group of movers initially accounts for 0.1×391.14 billion VKT. Their residential density increases by a factor of 3.69 from the initial level in region five, 861.14, to the final level in region nine, $1.1 \times 2,741.09$. This causes their driving to decrease by a factor of 0.90, for a total decrease of 3.91 billion VKT. Summing, this relocation causes a change in aggregate driving of $3.06 - 2.77 - 3.91 = -3.62$ billion VKT. Since aggregate driving is about 5.08 trillion VKT, this is a decrease of 0.07%.

To get a sense a sense of the costs of relocating 1% of the total MSA population or 2.5 million people, note that this policy ultimately requires the abandonment of about 1 million houses. At 200,000 dollars per unit, this is 200 billion dollars worth of housing. Using a 5% interest rate, the annualized value of this housing is about 10 billion dollars. Presumably, densification policies would allow housing to depreciate before being abandoned, so this we might expect this cost to be somewhat lower. On the other hand, with a 1 trillion dollar annual expenditure on road transportation, a 0.07% decline yields annual savings of 700 million dollars. Moreover, according to Parry et al. (2007), the external cost associated with car driving is about seven cents per kilometer.[37] Multiplying, the gain from less congestion, fewer accidents, and less pollution from a 0.07% reduction in driving is only 25 million dollars. Comparing these two results suggests that the value of reductions in driving is unlikely to be large relative to the costs of densification.

6.3. *Densification vs. gas taxes and congestion pricing*

Assessing the wisdom of using density changes to manage traffic requires that we evaluate the effects of urban form on driving, as we do above, and also that we compare density changes to other policies that we might use to manage driving, gasoline taxes and congestion pricing in particular.

There is a large literature on the relationship between gasoline prices and consumption. Hughes et al. (2015) survey this literature, while Coglianese et al. (2017) provide recent contributions to the literature. Because there may be many margins of adjustments in the long run following a change in gasoline prices, we rely price elasticities in the short run for which we expect gasoline consumption to reflect driving. While this elasticity appears to be about 0.3, there is evidence that it may have declined to 0.1 in the last decade (Hughes et al., 2015).

Using this very conservative estimate, a fifty percent increase in gasoline price causes a 5% reduction in total driving. This is about the same decrease in aggregate driving as was accomplished by the extreme relocation policy described above, but it is accomplished with price variation that is well within the range of prices experienced in the US between 2010 and 2015. If the objective of policy is to reduce aggregate driving, it is hard to imagine that gasoline taxes do not accomplish this objective at a lower cost than forcing density changes.

Congestion pricing schemes are also used as a tool to manage traffic in urban areas and involve time of day, area specific road tolls. The London congestion charge began in 2003 and required the payment of about 8 USD to enter central London, an area of about 22 square kilometers, during working hours. This policy led to a dramatic reduction in travel, about 34% for cars and 12% for all vehicle types, an increase in peak hour travel speeds from 14.3 to 16.7 km per hour, and a dramatic decrease in delay relative to free-flow travel speeds (Leape, 2006). The Singapore congestion charge began in 1975 and was about 2.5 USD per day. It converted from a paper-based to electronic enforcement system in 1999 with somewhat lower charges. At its beginning, this program was responsible for about a 45% reduction in peak area vehicle travel in the affected area and an increase in travel speeds from 19 to 36 km per hour (Santos, 2005). Stockholm is the third main city with a congestion pricing scheme. Begun in 2006, with a time of day charging that peaks at about 3 USD at peak hours and tapers to zero during off peak times, this program caused about 30% reduction in vehicles in the affected areas and a dramatic decrease in travel times (Borjesson et al., 2012).

Relative to the marginal and uncertain reductions in driving that appear to result from densification policies, it is hard to imagine that congestion pricing is not a more cost effective way to reduce urban congestion than is density changes.[38]

7. Conclusion

Urban density appears to have a small causal effect on driving. Our estimates of the density elasticity are generally between −7% and −10% and is about −8% in our preferred specification. The literature on this issue is large. Our estimates improve on those in the literature in four ways. First, we use better data. We are the first to use a data set as large as the NHTS to estimate the effect of urban form on driving using microdata describing households and their residential location. Second, we develop a parametric test for sorting. Although the literature has long been aware that cross-sectional differences in driving behavior across locations may reflect sorting, it has yet to develop a persuasive quasi-experimental design. Given this, our ability to test for sorting using cross-sectional travel survey data and panel about urban form is an advance. Third, we implement a quasi-experimental design for dealing with the possibility of endogenous determination of density. Specifically, we use subterranean geology to instrument for surface density. Fourth, our econometric model is motivated by a theoretical foundation. Ultimately, this means that we are able to recover the structural parameters governing the way that travel behavior responds to density. To the extent that we are able to check, these structural parameters appear to be consistent with related estimates in the literature. This structural model also highlights that, even if densification is welfare improving, it does not remove the need for congestion pricing. Whether neighborhoods are high density or low, without congestion pricing, drivers do not

[37] See also Santos et al. (2010) for further discussion of this issue. There is a obviously a range of values in the literature for the various components of the external costs associated with car driving (congestion, accidents, pollution, etc). The values taken by Parry et al. (2007) are on the high side because they disproportionately rely on US estimates that adopt the statistical value for human life used by the US Department of Transportation. The statistical value of human life is much higher in the US than in other developed countries.

[38] While congestion pricing appears to have dramatic effects on the volume and speed of travel, there is some debate over whether such programs are welfare improving. The central issue is that the demand for travel appears to be very elastic, so that deadweight loss from congestion is small, while the costs of implementing congestion pricing plans can be large. See Prud'homme and Bocarejo (2005) for a nice illustration of these issues, which are also discussed in Couture et al. (2018).

account for their contribution to congestion without an explicit pricing program.

Our estimates of the relationship of driving to urban form allow us to assess the cost effectiveness of densification as a policy response to excessive driving. These estimates suggest that urban form is not cost effective compared to explicit pricing programs. In particular, even concentrating the population residing in 83% of the area the continental US into an area of about 1500 square kilometers would result in only about a 5% decrease in aggregate driving, and this policy appears to describe the upper envelope of what densification policies can accomplish. On the other hand, existing estimates of the gasoline price elasticity of driving suggest that a similar decrease in driving would be accomplished with a gas tax that is no larger than gasoline price fluctuations observed over the past five to ten years. Congestion pricing programs appear to have even larger effects. In sum, while dense urban development may well be desirable because it provides a residential environment where people want to live and that allows them work more productively (e.g., Rosenthal and Strange, 2008), it is probably more costly to manipulate driving behavior through densification policies than through congestion pricing or gasoline taxes.

Appendix A. Generalization of sorting model

The econometric model of sorting developed in Section 3 assumes that the propensity to drive of immigrants to location j depends on the density of the region, but that emigrants are a representative random subset of current residents. We here generalize this model to allow the populations of both immigrants and emigrants to be systematically different from the population of current residents.

We maintain the same basic framework. Driving of a resident in location j is given by equation (11), the propensity to drive of an incumbent resident of location j is given by (12) and we continue to consider the movement of an exogenous share of residents, s. However, we now allow the propensity to drive of immigrants and emigrants to diverge and to depend on density. In particular, using an E superscript to denote emigrants and an I for immigrants, suppose that the propensity to drive for these two populations are

$$\theta^I = \zeta_0^I + \zeta_1^I x_j + \mu_{ij} \tag{A.1}$$

$$\theta^E = \zeta_0^E + \zeta_1^E x_j + \mu_{ij}, \tag{A.2}$$

and let $\Delta\zeta_k = \zeta_k^I - \zeta_k^E$. If share s of the population emigrates from j and is replaced according to this process, then we have the following analog to equation (14),

$$y_{ij}^1 = [(\alpha_0 + \alpha_1 x_j^0 + \mu_{ij}) + \beta x_j^1 + \delta_j] + s[\Delta\zeta_0 + \Delta\zeta_1 x_j^1] \tag{A.3}$$

$$= \alpha_0 + (\alpha_1 + \beta)x_j^0 + \delta s \Delta x_j + \beta \Delta x_j + \Delta\zeta_0 s + \Delta\zeta_1 \; s x_j^0 + \delta_j + \mu_{ij}. \tag{A.4}$$

All of the terms in this expression, except $\Delta\zeta_0 s$ and $\Delta\zeta_1 \; s x_j^0$ also appear in the corresponding expression in the main text, equation (14). The first is simply the level term in share of migrants. We include this term in our regressions anyhow. The second is a two way interaction. We include something very close to this term in some of our robustness checks, $\Delta\zeta_1 \; s x_j^1$ (column 3, Table 5).

Informally, this more general model of migration and sorting involves tripling the number of parameters that relate density to propensity to drive (from two to six). Not too surprisingly, when people migrate and density changes this leads to more interaction terms. This suggests that we should be cautious in our interpretation of the coefficients of the various interaction terms.

With this said, the basic intuition that motivates our approach appears robust. When people with different propensities to drive systematically choose different densities and density directly affects how much people drive, then we should expect that changes in density will have different effects than lagged levels.

Appendix B. Heckman estimations

In this appendix, we replicate the Heckman selection models estimated by much of prior research. We now estimate two equations. The first equation is Probit regression estimating the probability that a household lives in a high density area: $\text{Probit}(p_i) = x_{-den,i}\gamma + e_i$, where p_i is the probability of household i residing in a high-density area, x_{-den} is a set of explanatory variables as used in Table 2 which does not include density, and e is a normal error term with standard deviation σ_e. The second equation duplicates the VKT regressions estimated so far but also includes among its explanatory variables a transformation of the predicted probability to live at higher density, as estimated in the first equation: $y_i = x_i\beta + \sigma\lambda(x_{-den,i}\gamma/\sigma_e) + \epsilon_i$, where $\lambda(x_{-den,i}\gamma/\sigma_e)$ is computed as the inverse Mills ratio evaluated at $x_{-den,i}\gamma/\sigma_e$. Simply put, this approach treats the selection of households into high-density residential areas as a missing variable problem in the main regression and estimates this

Table 12
Heckman selection models (one-step maximum likelihood estimation).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	All	All	All	MSA	MSA	MSA	MSA	MSA
Selection into:	Above median density						Top density decile	
log 10-km density	−0.12[a]	−0.13[a]	−0.11[a]	−0.12[a]	−0.13[a]	−0.11[a]	−0.21[a]	−0.16[a]
	(0.0058)	(0.0064)	(0.010)	(0.0059)	(0.0064)	(0.010)	(0.022)	(0.026)
Controls:								
Demographics	Y	Y	Y	Y	Y	Y	Y	Y
Geography	N	Y	Y	N	Y	Y	Y	Y
Local socio-econ.	N	Y	Y	N	Y	Y	Y	Y
MSA fixed effects	N	N	Y	N	N	Y	N	Y
Observations	126,203	126,203	126,203	99,875	99,875	99,875	99,875	99,875

Notes: Results reported for the main regression using log household VKT as dependent variable. The selection equation regards above median MSA density in columns 1–6, and selection into the highest density decile in columns 6–8. The sample is all driving households in columns 1–3 and all MSA households in column 4–8. In columns 1–3, median density is defined relative the entire population of driving households whereas in columns 4–6, it is defined relative to driving households that live in MSAs. Robust standard errors in parentheses. [a], [b], [c]: significant at 1%, 5%, 10%. Demographic controls include a white/Asian indicator, log income, log household size, a single indicator, age, age squared, gender, education, and education squared. Geographic controls include average precipitation and its standard deviation and average temperature and its standard deviation. Local socio-economic controls include the share of residents with higher education and its square and log local income.

missing variable in a separate selection equation from the (non-linear) probability of residing in a high-density areas.

The results are reported in Table 12. In columns 1 to 6, we consider selection into neighborhoods with above median 10-kilometer density. Depending on the specification, we control for household demographics alone, add geography and socioeconomic controls, or also add MSA fixed effects. These first six columns of Table 12 estimate a density elasticity of VKT between −0.11 and −0.13. These elasticity are slightly larger in magnitude than those estimated in Table 2, but by only two or three percentage points. In columns 7 and 8, we consider a selection equation for a density threshold corresponding to the top decile of density for MSA households. The estimated density elasticity of VKT is now larger in magnitude, reaching −0.16 in a specification including MSA fixed effects. We further investigate the driving behavior at the top density decile in Section 5.6.

Appendix C. Robustness checks

Table 13
Robustness of selection estimations using local tenure length to measure mobility.

	(1) No high den. location	(2) No high VKT hh	(3) Big Δ & <50	(4) Small Δ & >60	(5) population density	(6) ind. day km as DV	(7) ind. day mn as DV	(8) speed as DV
log 10-km density 1990	−0.070^a	−0.060^a	−0.084^a	−0.054^a	−0.074^a	−0.12^a	−0.018^a	−0.10^a
	(0.0039)	(0.0055)	(0.0098)	(0.0088)	(0.0056)	(0.0079)	(0.0043)	(0.0049)
Δ_{90-10} log 10-km density	−0.029	−0.028	0.023	−0.030	−0.067^b	−0.15^a	0.00072	−0.15^a
	(0.030)	(0.033)	(0.063)	(0.45)	(0.026)	(0.041)	(0.031)	(0.023)
Mobility ×Δ log density	−0.00041	−0.00049	0.012^c	−0.0048	−0.0053	−0.019^a	−0.0050	−0.013^a
	(0.0038)	(0.0037)	(0.0068)	(0.040)	(0.0035)	(0.0047)	(0.0034)	(0.0030)
Mobility	−0.010^a	−0.0094^a	−0.0094	−0.024^c	−0.011^a	−0.0087^b	−0.0068^b	−0.00096
	(0.0038)	(0.0037)	(0.0068)	(0.040)	(0.0035)	(0.0047)	(0.0034)	(0.0030)
F-test 1 p-value	0.00022	0.017	0.21	0.75	0.000092	0	0.000017	0
F-test 2 p-value	0.17	0.31	0.082	0.96	0.76	0.39	0.52	0.042
R^2	0.37	0.34	0.25	0.27	0.37	0.18	0.12	0.14
Observations	74,864	90,662	18,711	19,979	99,875	83,313	85,996	82,849
Number of MSA	275	275	248	252	275	275	275	275

Notes: All regressions include MSA fixed effects. Robust standard errors clustered by MSA in parentheses. ^{a}, ^{b}, ^{c}: significant at 1%, 5%, 10%. The dependent variable and explanatory variables of interest are in log in all columns. Demographic controls include a white/Asian indicator, log income, log household size, a single indicator, age, age squared, gender, education, and education squared. Geographic controls include average precipitation and its standard deviation, and average temperature and its standard deviation. Local socio-economic controls include the share of residents with higher education and its square and log local income. F-test 1 is a joint test of the equality of the coefficients on initial log 10-km density and Δ log 10-km density and of the coefficient on Mobility ×Δ log density being zero. F-test 2 is a test of the equality of the coefficients on initial log 10-km density and Δ log 10-km density.

Table 14
Selection and mobility using information about the renter/homeowner status of the households.

Period Household sample	(1) 90 to 10 All	(2) 90 to 10 All	(3) 90 to 10 All	(4) 90 to 10 Big Δ	(5) 90 to 10 Small Δ	(6) 90 to 10 Age<50	(7) 00 to 10 All	(8) 00 to 10 All
Initial log 10-km density	−0.081^a	−0.080^a	−0.082^a	−0.085^a	−0.078^a	−0.084^a	−0.080^a	−0.080^a
	(0.0053)	(0.0052)	(0.0051)	(0.0067)	(0.0067)	(0.0072)	(0.0052)	(0.0050)
Δ log 10-km density	−0.057^a	−0.071^a	−0.074^a	−0.089^a	−0.081	−0.083^a	−0.053^a	−0.043^b
	(0.013)	(0.012)	(0.012)	(0.018)	(0.061)	(0.022)	(0.017)	(0.019)
Renter ×Δ log density	−0.20^a	0.013	0.040	−0.010	−0.087	0.040	0.033	0.032
	(0.030)	(0.036)	(0.038)	(0.046)	(0.16)	(0.035)	(0.050)	(0.050)
Renter		−0.15^a	−0.33^a	−0.14^a	−0.12^a	−0.14^a	−0.15^a	−0.15^a
		(0.013)	(0.094)	(0.020)	(0.040)	(0.018)	(0.011)	(0.012)
Renter × log density			0.015^b					
			(0.0073)					
Past Δ log 10-km density			–					−0.032
								(0.023)
F-test 1 p-value	0	0.60	0.37	0.95	0.84	0.49	0.10	0.13
F-test 2 p-value	0.026	0.38	0.44	0.80	0.97	0.93	0.078	0.020
R^2	0.37	0.37	0.37	0.36	0.37	0.26	0.37	0.37
Observations	99,875	99,875	99,875	46,942	48,939	39,253	99,875	99,875
Number of MSA	275	275	275	263	267	274	275	275

Notes: The dependent variables is log household VKT in all columns. All regressions are estimated with OLS and include MSA fixed effects with demographic controls (a white/Asian indicator, log income, log household size, a single indicator, age, age squared, gender, education, and education squared), geographic controls (average precipitation and its standard deviation, and average temperature and its standard deviation) and local socio-economic controls (the share of residents with higher education and its square and log local income). Robust standard errors clustered by MSA in parentheses. ^{a}, ^{b}, ^{c}: significant at 1%, 5%, 10%. F-test 1 is a joint test of the equality of the coefficients on Initial log 10-km density and Δ log 10-km density and of the coefficient on Mobility ×Δ log density being zero. F-test 2 is a test of the equality of the coefficients on initial log 10-km density and Δ log 10-km density.

Table 15
Robustness checks for sorting on demographics OLS estimations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Period	00 to 10	00 to 10	00 to 10	90 to 10	90 to 10	90 to 10	90 to 10	90 to 10
Household sample	All	<50	>60	All	All	Indiv.	All	All
Dependent var.:	an. km	an. km	an. km	stated km	odometer	ind. day km	an. km	an. km
Density:	10 km	10 km	10 km	10 km	10 km	10 km	1 km	NLCD 10 km
Initial log density	−0.082[a]	−0.087[a]	−0.075[a]	−0.12[a]	−0.094[a]	−0.14[a]	−0.053[a]	−0.040[a]
	(0.0053)	(0.0072)	(0.0054)	(0.0075)	(0.0060)	(0.0083)	(0.0037)	(0.0046)
Δ log density	−0.050[a]	−0.049[c]	−0.036	−0.066[a]	−0.077[a]	−0.066[b]	−0.047[a]	−0.039[a]
	(0.017)	(0.028)	(0.035)	(0.024)	(0.024)	(0.029)	(0.0060)	(0.0069)
Past Δ density		−0.037	−0.015					
		(0.028)	(0.039)					
Controls:								
Demographics	Y	Y	Y	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y	Y	Y	Y
Local socio-econ.	Y	Y	Y	Y	Y	Y	Y	Y
Decade indicators	N	N	N	Y	Y	Y	Y	Y
Decade × log density	N	N	N	Y	Y	Y	Y	Y
Decade ×Δ log density	N	N	N	Y	Y	Y	Y	Y
F-test 1 p-value	.	.	.	0.0001	0.027	0	0.0034	0.0033
F-test 2 p-value	0.039	0.10	0.27	0.015	0.40	0.0062	0.22	0.91
R^2	0.37	0.26	0.26	0.42	0.43	0.09	0.37	0.37
Observations	99,875	39,253	40,421	93,602	71,742	121,808	99,874	99,423
Number of MSA	275	274	274	275	275	275	275	275

Notes: All regressions include MSA fixed effects. Robust standard errors clustered by MSA in parentheses. [a], [b], [c]: significant at 1%, 5%, 10%. The dependent variables and explanatory variables of interest are in log in all columns. Demographic controls include a white/Asian indicator, log income, log household size, a single indicator, age, age squared, gender, education, and education squared. Geographic controls include average precipitation and its standard deviation, and average temperature and its standard deviation. Local socio-economic controls include the share of residents with higher education and its square and log local income. When decade effects are introduced, households in their 40s are used as reference. F-test 1 is a joint test of the equality of the coefficients on initial log 10-km density and Δ log 10-km density and of the coefficients on decade indicators interacted with Δ log density all being zero. F-test 2 is a test of the equality of the coefficients on initial log 10-km density and Δ log 10-km density.

References

Ahlfeldt, G.M., Redding, S.J., Sturm, D.M., Wolf, N., 2015. The economics of density: evidence from the Berlin wall. Econometrica 83 (6), 2127–2189.
Baum-Snow, N., 2007. Did highways cause suburbanization? Quart. J. Econ. 122 (2), 775–805.
Bento, A., Cropper, M.L., Mobarak, A.M., Vinha, K., 2005. The effects of urban spatial structure on travel demand in the United States. Rev. Econ. Stat. 87 (3), 466–478.
Blaudin de Thé, C., Lafourcade, M., 2015. The carbon footprint of suburbanization: Evidence from French household data. Processed, Paris School of Economics.
Boarnet, M.G., 2011. A broader context for land use and travel behavior, and a research agenda. J. Am. Plan. Assoc. 77 (3), 197–213.
Borjesson, M., Eliasson, J., Hugosson, M.B., Brundell-Freij, K., 2012. The stockholm congestion charges-5 years on. effects, acceptability and lessons learnt. Transp. Policy (Oxf) 20 (1), 1–12.
Brownstone, D., Golob, T.F., 2009. The impact of residential density on vehicle usage and energy consumption. J. Urban Econ. 65 (1), 91–98.
Burchfield, M., Overman, H.G., Puga, D., Turner, M.A., 2006. Causes of sprawl: a portrait from space. Quart. J. Econ. 121 (2), 587–634.
Cao, X., Mokhtarian, P.L., Handy, S.L., 2009. Examining the impacts of residential self-selection on travel behaviour: a focus on empirical findings. Trans. Rev. 29 (3), 359–395.
Cervero, R., Kockelman, K., 1997. Travel demand and the 3Ds: density, diversity, and design. Transport. Res. Part D Transp.Environ. 2 (3), 199–219.
Coglianese, J., Davis, L.W., Kilian, L., Stock, J.H., 2017. Anticipation, tax avoidance, and the price elasticity of gasoline demand. J. Appl. Economet. 32 (1), 1–15.
Combes, P.-P., Gobillon, L., 2015. The Empirics of Agglomeration Economies. In: Duranton, G., Henderson, V., Strange, W. (Eds.), Handbook of Regional and Urban Economics, 5A. Elsevier, Amsterdam, pp. 247–348.
Couture, V., 2014. Valuing the consumption benefits of urban density. Processed, University of California Berkeley.
Couture, V., Duranton, G., Turner, M.A., 2018. Speed. Rev. Econ. Stat. 100 (4), 725–739.
Duranton, G., Turner, M.A., 2011. The fundamental law of road congestion: evidence from US cities. Am. Econ. Rev. 101 (6), 2616–2652.
Environmental Systems Research Institute, 1998a. ESRI Data and Maps 1998: Disc 2 - Zip Code Tabulation Areas. This map data was included with an ArcGIS 7 License.
Environmental Systems Research Institute, 1998b. ESRI Data and Maps 1998: Discs 3&4 - Block Group Boundaries. This map data was included with an ArcGIS 7 License.
Environmental Systems Research Institute, 2004. ESRI Data and Maps 2004: United States Disc - Block Group Boundaries. This map data was included with an ArcGIS 9 Licence.
Ewing, R., Cervero, R., 2001. Travel and the built environment: a synthesis. Transport. Res. Rec. J.Transport. Res. Board 0 (1780), 87–114.
Ewing, R., Cervero, R., 2010. Travel and the built environment: a meta-analysis. J. Am. Plan. Assoc. 76 (3), 265–294.
Federal Highway Administration, 2005. National highway planning network. http://www.fhwa.dot.gov. Accessed: 2014-05-24.
Giuliano, G., Small, K.A., 1993. Is the journey to work explained by urban structure? Urban Studies 30 (9), 1485–1500.
Glaeser, E.L., Kahn, M.E., 2004. Sprawl and Urban Growth. In: Henderson, V., Thisse, J.-F. (Eds.), Handbook of Regional and Urban Economics, 4. North-Holland, Amsterdam, pp. 2481–2527.
Glaeser, E.L., Kahn, M.E., 2008. The greenness of cities: carbon dioxide emissions and urban development. J. Urban Econ. 67 (3), 404–418.
Gordon, P., Kumar, A., Richardson, H.W., 1989. The influence of metropolitan spatial structure on commuting time. J. Urban Econ. 26 (2), 138–151.
Handy, S., 2005. Smart growth and the transportation-land use connection: what does the research tell us? Int. Reg. Sci. Rev. 28 (2), 146–167.
Head, K., Mayer, T., 2014. Gravity Equations: Workhorse, Toolkit, and Cookbook. In: Gopinath, G., Helpman, E., Rogoff, K. (Eds.), Handbook of International Economics, 4. North-Holland, Amsterdam, pp. 131–195.
Hughes, J., Knittel, C., Sperling, D., 2015. Evidence of a shift in the short-run price elasticity of gasoline demand. Energ. J. 29 (1), 113–134.
Intergovernmental Panel on Climate Change, 2007. Working group III contribution to fouth assessment report of the IPCC: Mitigation of climate change. Cambridge University Press, Cambridge, UK and New York, NY, USA.
Intergovernmental Panel on Climate Change, 2014. Working group III contribution to fifth assessment report of the IPCC: Mitigation of climate change. Cambridge University Press, NY, New York, USA.
Kuzmyak, J.R., 2012. Land use and traffic congestion. Professional Paper, 618. Arizona DOT and USDOT.
Leape, J., 2006. The London congestion charge. J. Econ. Perspect. 20, 157–176.
Missouri Census Data Center, 1990. Block-group and county populations in 1990. http://www.mcdc.missouri.edu/websas/geocorr90.shtml. Accessed: 2014-01-14.
Missouri Census Data Center, 2000. Block-group and county populations in 2000. http://www.mcdc.missouri.edu/websas/geocorr2k.html. Accessed: 2014-01-14.
Missouri Census Data Center, 2010. Block-group and county populations in 2010. http://www.mcdc.missouri.edu/websas/geocorr12.shtml. Accessed: 2014-01-14.
National Historical Geographic Information System, 2010. Block-group level demographic variables for 2006–2010 and 2007–2011. http://www.nhgis.org. Accessed: 2014-01-30.
Oster, E., 2017. Unobservable selection and coefficient stability: theory and validation. J. Busin. Econ. Stat. 0 (0), 1–18.
Parry, I.W., Walls, M., Harrington, W., 2007. Automobile externalities and policies. J. Econ. Lit. 45 (2), 1335–1353.
PRISM Climate Group at Oregon State University, 2012a. United States Average January Mean Temperature, 1981–2010. http://www.prism.oregonstate.edu. Accessed: 2014-04-23.

PRISM Climate Group at Oregon State University, 2012b. United States Average January Precipitation, 1981–2010. http://www.prism.oregonstate.edu. Accessed: 2014-04-23.

Prud'homme, R., Bocarejo, J., 2005. The London congestion charge: a tentative economic appraisal. Transp. Policy (Oxf) 12 (3), 279–287.

Rosenthal, S.S., Strange, W.C., 2008. The attenuation of human capital spillovers. J. Urban Econ. 64 (2), 373–389.

Santos, G., 2005. Urban congestion charging: a comparison between London and Singapore. Transp. Rev. 25 (5), 511–534.

Santos, G., Behrendt, H., Maconi, L., Shirvani, T., Teytelboym, A., 2010. Part i: externalities and economic policies in road transport. Res. Transport. Econ. 28 (1), 2–45.

Small, K.A., Verhoef, E.T., 2007. The economics of urban transportation. Routledge, New York (NY).

Stevens, M.R., 2017. Does compact development make people drive less? J. Am. Plan. Assoc. 83 (1), 7–18.

United States Geological Survey, 2000. NLCD 1992 Land Cover. http://www.mrlc.gov. Accessed: 2013-12-19.

United States Geological Survey, 2001. Landslide Incidence and Susceptibility in the Conterminous United States. http://www.nationalatlas.gov/atlasftp.html. Accessed: 2014-05-20.

United States Geological Survey, 2003. Principal Aquifers of the 48 Conterminous United States. http://www.nationalatlas.gov/atlasftp.html. Accessed: 2014-05-20.

United States Geological Survey, 2005. Significant United States Earthquakes, 1568–2004. http://www.nationalatlas.gov/atlasftp.html. Accessed: 2014-05-20.

United States Geological Survey, 2011a. NLCD 2006 Land Cover Version 2.0. http://www.mrlc.gov. Accessed: 2014-01-24.

United States Geological Survey, 2011b. NLCD 2001 land cover version 2.0. http://www.mrlc.gov. Accessed: 2014-01-24.

U.S. Department of Commerce, U.S. Census Bureau, Geography Division, 2010. Block-group level US Census 2010 geography Shapefiles.

U.S. Department of Commerce, U.S. Census Bureau, Geography Division, 2010. Tiger/line shapefile, 2010, 2010 nation, u.s., 2010 census 5-digit zip code tabulation area (zcta5) national. http://www.census.gov/geo/www/tiger. Accessed: 2014-09-06.

U.S. Department of Transportation, Federal Highway Administration, 2009. 2009 National Household Travel Survey. http://www.nhts.ornl.gov. Accessed: 2014-06-13.