Traffic Fatalities and Population Density

Nils-Petter Lagerlöf[†]

September 2023

Abstract

What are the cross-sectional correlates and/or causes of traffic safety? Here I document a negative association between traffic fatality rates and population density across U.S. counties and European regions. This is robust to various controls and state/country fixed effects. For the U.S., it also holds when replacing, or instrumenting, modern population density with density in 1900, prior to the spread of cars, possibly suggesting a causal link. Traffic fatality rates also show a negative correlation with per-capita incomes, but less robustly than with population density. I also find a positive association between homicides and traffic deaths, suggesting a connection to other forms of violence. A simple model is presented to help interpret some of these results.

[†]Department of Economics, York University. E-mail: lagerlof@yorku.ca.

1 Introduction

Which are the major correlates, and possible causes, of traffic deaths? Economists have studied how traffic safety responds to specific events, technologies, and policies, such as the COVID-19 pandemic, changes in vehicle sizes, and congestion charges. (See Section 2 below.) However, no one has yet systematically explored how locations with high rates of traffic fatalities differ from other locations.

In this paper, I document a robust negative correlation between population density and traffic fatality rates across both U.S. counties and European regions. At one level, this may seem unsurprising, since higher population density is linked to lower dependence on cars. However, it is still useful to have documented, not least because cities are known to be more adversely affected by other traffic externalities, such as congestion and air pollution. Traffic fatality rates show the opposite pattern: they are lower in more densely populated places.

It is also arguably surprising how robust this result is. The negative association between population density and traffic fatality rates holds when controlling for per-capita incomes and life expectancy. It holds with fixed effects for U.S. states, and European countries, respectively, and within the largest U.S. states and European countries when examined on their own. It also holds when using urbanization rates instead of population density.

Although the U.S. and European data are not completely comparable (e.g., referring to different years), comparisons suggest that differences in population density between the U.S. and Europe can partly, but not fully, explain differences in traffic fatalities.

Meanwhile, there is no equally robust correlation between per-capita incomes and traffic fatalities, although the association is mostly negative when significant: richer locations tend to have lower traffic fatality rates.

Controlling for homicides does not change the estimated correlations between traffic deaths and density, so these findings do not reflect any general tendency for urban locations to be less violent than rural ones. (Indeed, the opposite is true for Europe.) However, higher homicide rates tend to be associated with higher traffic fatality rates, conditional on density and other factors. This finding suggests that traffic deaths may be partly linked to more general forms of violence.

The results for the U.S. hold when replacing, or instrumenting, current population density with the same measure in 1900, prior to the spread of cars. This may suggest a causal relationship running from population density to traffic fatalities.

I also explore some alternative outcome variables, sometimes used in the existing literature, in particular vehicles per capita and the fraction commuters using public transit. These show similar patterns but not as clearly as traffic fatality rates do. The explanatory power of population density and the other controls tend to be lower. My suggested interpretation is that traffic deaths better capture variation in the extent of overall car dependence.

Informed by these empirical patterns, I propose a simple model that may help us think about the political-economy mechanisms driving them. Agents choose mode of transportation, and vote on taxes for public transit. Once a location's population density becomes high enough, it switches from a regime with no taxes and only driving to one with some taxation and some agents using public transit, although drivers are still in majority. Further increases in density at some point cause a switch to a regime where non-drivers become a majority, making taxes and public transit provision increase further.

The model is able to reproduce the empirical pattern that higher population density is associated with a larger total number of drivers (i.e., more congestion), but a lower fraction of the population driving, and thus lower traffic fatality rates.

By contrast, higher per-capita incomes may in this model be associated with higher or lower traffic fatality rates, which matches the finding that population density is more robustly correlated with traffic fatality rates than per-capita incomes are.

The rest of this paper is organized as follows. Next, Section 2 relates this paper to some of the existing literature. Section 3 provides a background by showing 20th-century traffic fatality trends for the United States as a whole. Section 4 sets up the model. The compilation of the datasets used in the empirical analysis is explained in Section 5. Section 6 presents the empirical results. Section 7 concludes.

2 Existing Literature

This paper touches on a vast literature on density/urbanization, congestion, and transportation (see, e.g., Duranton et al., 2011; Anderson, 2014; Duranton and Turner, 2018; Seidel and Wickerath, 2020; Basso et al., 2021; Akbar et al., 2023). These papers mostly center on questions about congestion externalities, and ways to address these, e.g., through public transit investment. The focus is less on traffic deaths.

Some papers document negative effects on urban population density from exogenous factors related to cars, such as highways and car manufacturing (Baum-Snow, 2007, 2010; Ostermeijer et al., 2022). My interpretation of the correlations that I document is that they rather (at least in part) reflect a causal link from density to car dependence, specifically traffic fatalities, since they hold also when measuring density in 1900. More generally, there is a strong and well documented spatial persistence of urban agglomeration. However, this does not rule out effects running in the opposite direction as well, at least in some contexts.

The observation that higher population density is associated with lower car dependence is well known and has been documented and discussed by, e.g., Newman and Kenworthy (1989), Kenworthy and Laube (1999), and McIntosh et al. (2014). This body of research typically compares different large cities worldwide, rather than rural and urban locations in the same country. Also, they do not study traffic deaths.

There are a few papers on how road safety depends on congestion and/or population density. The introduction of congestion charges in London, and the associated decline in congestion, seems to have reduced collisions overall, with more mixed results for fatalities (Green et al., 2016; Tang and Van Ommeren, 2022). Others study the decline in congestion following the COVID-19 pandemic, also with mixed results on fatalities (Yasin et al., 2021; Hughes et al., 2023).[1] However, none of these contributions systematically explores the relationship between traffic fatalities and population density in cross-sectional data.

The positive correlations that I find between homicides and traffic fatalities relate indi-

[1]See also Retallack and Ostendorf (2019) and Albalate and Fageda (2021) for other studies finding mixed and/or non-monotonic relationships between congestion and fatalities.

rectly to research by Beland and Brent (2018) about the effects on domestic violence from stress due to congestion. That type of link does not seem likely to drive the results I find, since more congestion (or higher density) is associated with lower traffic fatality rates in my data.

Research on traffic fatalities often take an interest in vehicle sizes (Anderson, 2008; Jacobsen, 2013; Anderson and Auffhammer, 2014; Tyndall, 2021). This may be one factor driving the urban-rural relationship between density and traffic fatalities: as pointed out by Jacobsen (2013, p. 20), "drivers who currently choose large vehicles tend to live in rural areas." Many papers in this literature are interested in safety conditional on the amount of driving, or being involved in a crash, which leaves out car dependence itself as a causal factor. (Tyndall, 2021, is an exception.)

Other recent work on traffic safety explores racial disparities (Chalfin and Massenkoff, 2022), safety in informal public transit networks (Schönholzer et al., 2022), and the effects of cell phone bans (Wright and Dorilas, 2022) and speed-limit campaigns (Bauernschuster and Rekers, 2022).

The correlations that I document between per-capita incomes (or per-capita GDP) and traffic fatalities relate to research on the links between income shocks and traffic fatalities (Ruhm, 1995, 2000; Maheshri and Winston, 2016; French and Gumus, 2021). This research tends to find a decline in traffic fatalities during recessions, implying that lower incomes are associated with safer roads. This contrasts with my findings in cross-sectional data of a negative relationship between traffic fatalities and per-capita incomes. However, as noted, those correlations tend to be less robust than the ones referring to density.

There is also a large theoretical literature on public transit subsidies and how these are determined, politically and/or optimally (see, e.g., Vickrey, 1980; Parry and Small, 2005, 2009; De Borger and Proost, 2012, 2015). The model presented here differs in its direct focus on the effects of changes to population densities and incomes.

3 Background: Aggregate Time Trends in the United States

The focus of this paper is on the spatial variation in traffic fatalities, population density, and other variables. However, it is useful to first look at changes over time since cars started to spread in the early 20th century.

Figure 1 shows aggregate data on car (or vehicle) ownership and traffic fatalities for the United States since 1913. In the early 20th century there were very few cars and low traffic fatality rates. With the spread of the car deaths rose rapidly, peaking before WWII. Starting in the 1950s a number of policies were introduced to address deaths and injuries on U.S. roads, e.g., speed limits and restrictions on driving while intoxicated. Since the 1960s traffic fatality rates have declined sharply. Meanwhile, car ownership has continued to increase, but at a slower pace.

Fatalities have plateaued from around the 2010s, and started to rise again in recent years. This has been partly attributed to an increase in the size of vehicles (Tyndall, 2021).

We may discern some short-run co-movements in car ownership and traffic fatalities in Figure 1 (cf. the discussion in Section 2). For example, both seem to contract slightly around the 2008-2009 recession. This is easier to see when looking from 1990; see Figure A.1 in the Online Appendix.

To get a quantitative gauge, I regressed log per-capita traffic fatalities on log per-capita car ownership across these 109 years, including a squared time trend as control to help account for some of the confounding factors (such as long-term changes in technology and incomes). The point estimate of the coefficient on log car ownership comes to around .73 (and is significant at the 1% level with robust standard errors). That is, a 1% increase in per-capita car ownership is associated with a .73% increase in per-capita traffic fatalities. If we restrict the data set to the period from 1966 (a local peak in traffic fatalities), using the same specification, the same elasticity comes to almost .90%.

Although aggregate data can only tell us so much, this may help motivate the interpretation of traffic fatalities as a broad indicator of how much a society relies and depends on cars.

4 A Model

This section presents a framework for thinking about spatial variation in per-capita incomes, population density, and driving. Although the empirical analysis will mostly use data on traffic fatality rates, I do not model crashes between vehicles explicitly, but rather postulate a relationship between total traffic fatalities and the total number of agents who drive (see Section 4.3.3 below).

Let an agent j living in location i care about consumption (C_{ij}) and a good that I call just transportation (T_{ij}). Preferences are described by

$$V_{ij} = \alpha \ln C_{ij} + (1 - \alpha) \ln T_{ij}, \tag{1}$$

where $\alpha \in (0, 1)$. Thus, rather than modelling transportation as, e.g., time spent commuting to work I simply put it into the utility function.

The agent's transportation consumption can take one of two binary forms: driving or using public transit. Driving generates a transportation consumption of $T_{ij} = \eta_j X/(z_i L_i)$, where X is an exogenous amount of space (or roads) on which agents can drive, which is the same across locations, L_i is the population in location i, z_i is the fraction of that population who drives, and $\eta_j \geq 1$ is agent j's preference parameter for driving.

We may interpret driving as incorporating life-style and housing choices (e.g., suburban vs. downtown living), such that η_j is a life-style preference parameter. An agent with $\eta_j = 1$ is the most keen to stop driving; an agent with a high η_j is very averse to a non-car lifestyle. Alternatively, a low η_j could mean that the agent is worse affected by congestion when driving (cf. Anderson, 2014).

I assume that η_j is Pareto distributed, with a cumulative distribution function given by

$$F(\eta_j) = 1 - \eta_j^{-\frac{1}{\delta}}, \tag{2}$$

where $\delta \in (0, 1)$. As $\delta \to 0$ heterogeneity goes away, and all agents become idential with $\eta_j = 1$ for all j. Assuming $\delta < 1$ ensures that a mean exists for η_j.

Public transportation is a tax-funded non-rivalrous public good, which generates transportation consumption equal to $T_{ij} = P_i$ for all agents in location i who choose not to drive. I return below to how P_i is determined.

More compactly, transportation consumption can now be written:

$$T_{ij} = \begin{cases} \frac{\eta_j X}{z_i L_i} & \text{if agent } j \text{ drives,} \\ P_i & \text{if agent } j \text{ does not drive.} \end{cases} \tag{3}$$

Here it may be helpful to take stock and reflect on some of the braver model assumptions so far. Letting public transit be a non-rivalrous public good captures the idea that buses and trains use less space per passenger than cars do. Allowing for some congestion in public transit, as with driving, would not alter the results qualitatively, as long as this congestion effect is weaker than it is for driving.[2] Also, a public transit passenger may care about many other factors than congestion, such as network size, fleet capacity, and train/bus frequency, which are easier to interpret as non-rivalrous public goods.

The assumption that road space is exogenously given is meant to capture that roads require less public spending than public transit does (although driving does incur a private cost; see below). This assumption is done mostly for simplicity; road space does not need to come at exactly zero cost for the results to hold.

4.1 Income, Taxes, and Consumption

Let y_i denote pre-tax income per agent in location i, and τ_i the public-transit tax. (I abstract from other taxes.) Letting D denote the exogenous private cost of driving (capturing insurance, repairs, the car itself, etc.), the budget constraint can be written

$$C_{ij} = \begin{cases} (1-\tau_i)y_i - D & \text{if agent } j \text{ drives,} \\ (1-\tau_i)y_i & \text{if agent } j \text{ does not drive.} \end{cases} \tag{4}$$

It is assumed throughout that $y_i \geq D/\alpha$, which does not matter for the results qualitatively, but reduces the number of cases we need to consider. (It ensures that some agents will always drive in an equilibrium with public transit taxes.)

Incomes do not differ between agents in a location, so all agent heterogeneity is captured by the preference parameter η_j. One way to introduce an agent-specific income

[2] Specifically, I could let public transit users have a transportation consumption of $T_{ij} = P_i/[(1-z_i)L_i]^{\kappa}$ for some $\kappa \in [0,1)$. This would not change the results qualitatively. Here I consider the case where $\kappa = 0$.

component without changing any results would be to let both the driving cost and incomes vary proportionally between agents, leaving the driver/non-driver consumption ratio the same across agents.[3]

4.2 Public Transit Provision

Provision of public transit depends on aggregate tax revenue. Taxes are levied uniformly, with each of the L_i agents paying $\tau_i y_i$ in taxes, so that total tax revenue equals $\tau_i y_i L_i$. This revenue is transformed into public transit according to the production function

$$P_i = (\tau_i y_i L_i)^{\lambda}, \tag{5}$$

where we assume that $\lambda \in (0,1)$, so that the marginal effect of resources spent on transit declines with the amount spent.

4.3 Equilibrium

The equilibrium in location i is such that (i) agents choose to drive, or not drive, to maximize individual utilities; and (ii) the (uniform) tax rate, τ_i, maximizes utility of the median voter. Part (ii) captures the model's political-economy mechanism. Part (i) means we can substitute (3), (4) and (5) into (1), to write the utility of an agent in location i as

$$V_{ij} = \max\left\{V_{ij}^{D}, V_{ij}^{ND}\right\}, \tag{6}$$

where V_{ij}^{D} and V_{ij}^{ND} are utilities associated with driving and not driving, respectively, defined as follows:

$$V_{ij}^{D} = \alpha \ln\left[(1-\tau_i)y_i - D\right] + (1-\alpha)\ln\left(\frac{\eta_j X}{z_i L_i}\right), \tag{7}$$

and

$$V_{ij}^{ND} = \alpha \ln\left[(1-\tau_i)y_i\right] + (1-\alpha)\lambda \ln\left(\tau_i y_i L_i\right). \tag{8}$$

[3] That is, suppose that an agent j in location i earns income $y_{ij} = A_i h_j$ and faces a driving cost of $h_j D$. Then the ratio $[(1-\tau_i)A_i h_j - h_j D]/[(1-\tau_i)A_i h_j]$ boils down to the same as if using (4), but with A_i replacing y_i.

4.3.1 Tax Rates

Section A.1 in the Online Appendix derives the public transit tax and the fraction driving in equilibrium. There are three types of equilibrium and changing L_i shifts the economy from one equilibrium to another at certain thresholds.

First, an equilibrium where all agents choose to drive ($z_i = 1$) must be such that taxes are zero ($\tau_i = 0$), since there is no point funding transit if no one uses it.

Next, let τ_i^* be the tax rate that drivers prefer if some other agents are not driving. This will be the tax rate in an equilibrium where drivers make up a majority, but not 100% of the population, i.e., $z_i \in (1/2, 1)$. Intuitively, improved public transit eases congestion, so drivers may prefer a positive tax rate.[4]

As shown in Section A.1.2 of the Online Appendix, τ_i^* is defined from $G(\tau_i^*) \equiv 0$, where

$$G(\tau) = -\frac{\alpha}{1-\tau} - \frac{\delta \alpha y_i}{(1-\tau)\, y_i - D} + \frac{\lambda(1-\alpha)}{\tau}. \tag{9}$$

This implicitly defines $\tau_i^* > 0$ in terms of exogenous variables other than L_i.

Similarly, let

$$\tau_i^{**} = \frac{\lambda(1-\alpha)}{\alpha + \lambda(1-\alpha)} \tag{10}$$

be the tax rate preferred by those who choose to use public transit. This is the equilibrium tax rate if non-drivers are a majority, $z_i < 1/2$, and maximizes V_{ij}^{ND} in (8).

4.3.2 Population Thresholds

The type of equilibrium depends on population, L_i. Let $\widehat{L}_i$ denote the level of L_i below which the equilibrium is such that all agents drive. I show in the appendix that we can write this as

$$\widehat{L}_i = \left(\frac{B_i^*}{z_i^*}\right)^{\frac{1+\delta}{1+\lambda}}, \tag{11}$$

where

$$B_i^* = \left[\frac{(1-\tau_i^*)y_i - D}{(1-\tau_i^*)y_i}\right]^{\frac{\alpha}{(1-\alpha)(1+\delta)}} \left[\frac{X}{(\tau_i^* y_i)^{\lambda}}\right]^{\frac{1}{1+\delta}}, \tag{12}$$

[4] For examples and discussion of such political preferences, see, e.g., Anderson (2014).

and

$$z_i^* = \left[\frac{(1-\tau_i^*)y_i - D}{y_i - D} \right]^{\frac{\alpha}{1-\alpha}} < 1, \tag{13}$$

where the inequality comes from $\tau_i^* > 0$. We can interpret z_i^* as the level to which z_i must fall to induce those agents who still drive to support a tax of τ_i^*, which (recall) is leveled uniformly on both drivers and non-drivers. The factor B_i^* contains variables, other than L_i, that determine z_i; see (18) below.

Next, let $\widehat{\widehat{L}}_i$ denote the level of L_i above which the equilibrium is one with non-drivers in majority, $z_i < 1/2$. I show in Section A.1.5 of the Online Appendix that

$$\widehat{\widehat{L}}_i = (2B_i^{**})^{\frac{1+\delta}{1+\lambda}}, \tag{14}$$

where B_i^{**} is defined similarly as B_i^* in (12), except that τ_i^* is replaced by τ_i^{**}:

$$B_i^{**} = \left[\frac{(1-\tau_i^{**})y_i - D}{(1-\tau_i^{**})y_i} \right]^{\frac{\alpha}{(1-\alpha)(1+\delta)}} \left[\frac{X}{(\tau_i^{**} y_i)^{\lambda}} \right]^{\frac{1}{1+\delta}}. \tag{15}$$

The intuition behind how $\widehat{L}_i$ and $\widehat{\widehat{L}}_i$ are derived, and what B_i^{**} and B_i^* mean, will become more clear soon.

Having defined the population thresholds $\widehat{L}_i$ and $\widehat{\widehat{L}}_i$, we can now write the equilibrium tax rate as

$$\tau_i = \begin{cases} 0 & \text{if } L_i < \widehat{L}_i, \\ \tau_i^* & \text{if } L_i \in [\widehat{L}_i, \widehat{\widehat{L}}_i), \\ \tau_i^{**} & \text{if } L_i \geq \widehat{\widehat{L}}_i, \end{cases} \tag{16}$$

and the equilibrium fraction driving as

$$z_i = \begin{cases} 1 & \text{if } L_i < \widehat{L}_i, \\ Z(y_i, \tau_i^*, L_i) & \text{if } L_i \in [\widehat{L}_i, \widehat{\widehat{L}}_i), \\ Z(y_i, \tau_i^{**}, L_i) & \text{if } L_i \geq \widehat{\widehat{L}}_i, \end{cases} \tag{17}$$

where

$$Z(y_i, \tau_i, L_i) = \left[\frac{(1-\tau_i)y_i - D}{(1-\tau_i)y_i}\right]^{\frac{\alpha}{(1-\alpha)(1+\delta)}} \left[\frac{X}{(\tau_i y_i)^{\lambda}}\right]^{\frac{1}{1+\delta}} L_i^{-\frac{1+\lambda}{1+\delta}}. \tag{18}$$

Note from (12) and (18) that $B_i^* = Z(y_i, \tau_i^*, 1)$, i.e., the first two factors on the right-hand side in (18) evaluated at $\tau_i = \tau_i^*$. It can now be seen from (11), (12), (17), and (18) that a population level of $\widehat{L}_i$, and a tax rate of τ_i^*, generate a fraction driving of $Z(y_i, \tau_i^*, \widehat{L}_i) = z_i^*$, where z_i^* is given by (13).[5]

Analogously, to understand where $\widehat{\widehat{L}}_i$ comes from, note from (14), (15), (17), and (18) that $Z(y_i, \tau_i^{**}, \widehat{\widehat{L}}_i) = 1/2$. That is, $\widehat{\widehat{L}}_i$ is the population level where non-drivers can form a majority and implement their preferred tax rate, τ_i^{**}.

From (17) and (18) we can easily find the total number of drivers, $z_i L_i$.

4.3.3 Traffic Fatalities

The last step in the model analysis is to introduce some metric of traffic fatality rates. As discussed already, the model structure does not lend itself easily to explicitly capture collisions. To bring us a little closer to interpreting traffic fatalities, I simply assume an exogenous relationship between total traffic deaths, denoted K_i, and two other factors that the model can capture: (1) the total number of drivers, $z_i L_i$, and (2) speed of driving, which I let be proxied by space per driver, $X/(z_i L_i)$. I use this functional form:

$$K_i = (z_i L_i)^{\gamma} \left(\frac{X}{z_i L_i}\right)^{\beta}, \tag{19}$$

where $\gamma > 0$ and $\beta > 0$, i.e., both higher speeds and more drivers lead to more total fatalities. It also makes sense to assume that $\gamma > \beta$, meaning the net effect on total traffic fatalities from an increase in the total number of drivers is positive.[6]

It now follows that the traffic fatality rate, K_i/L_i, depends on total population according to

$$\frac{K_i}{L_i} = X^{\beta} z_i^{\gamma-\beta} L_i^{\gamma-\beta-1}. \tag{20}$$

[5] To see it another way, use (17) and (18) to note that $z_i^* = Z(y_i, \tau_i^*, \widehat{L}_i) = Z(y_i, \tau_i^*, 1)\widehat{L}_i^{-\frac{1+\lambda}{1+\delta}} = B_i^* \widehat{L}_i^{-\frac{1+\lambda}{1+\delta}}$. Solving for $\widehat{L}_i$ gives (11).

[6] The theoretical link from more congestion to lower speeds, and thus lower fatalities, has been noted by, e.g., Shefer and Rietveld (1997), Green et al. (2016), and Tang and Van Ommeren (2022).

If $\gamma = 1 + \beta$, then the traffic fatality rate is independent of population (L_i), and all changes are driven by the fraction who drive (z_i), meaning the traffic fatality rate is constant when all drive ($z_i = 1$). Setting γ slightly below $1 + \beta$ allows the fatality rate to be strictly decreasing in population, also when $z_i = 1$, but the more interesting results refer to the case when $z_i < 1$ (see below).

4.3.4 Numerical Illustration

Figure 2 shows how the equilibrium outcomes change with L_i for one numerical example ($\alpha = .25$; $y_i = 10$; $X = 1$; $D = 1.25$; $\beta = 1$; $\gamma = 1.95$; $\lambda = .2$; $\delta = .8$). These numbers are arbitrary, but the qualitative features hold under a broad set of parameter values. The most important assumptions here are that $\delta > \lambda$ and $\gamma < 1 + \beta$, as discussed below.

I then vary L_i from close to zero to above $\hat{\hat{L}}_i$. The top panel in Figure 2 shows how the tax rate makes discrete jumps when population reaches the two thresholds $\hat{L}_i$ and $\hat{\hat{L}}_i$, and the panel right below shows the associated shifts in the fraction of the population who drive.

The total number of drivers (shown in the third panel from the top) also shifts down when population reaches these thresholds, but is otherwise increasing, which hinges on $\delta > \lambda$, since $z_i L_i$ is proportional to $L_i^{\frac{\delta-\lambda}{1+\delta}}$; cf. (18). Intuitively, public transit must be costly enough, and driving valued sufficiently by the marginal driver, for a larger population to be associated with more drivers.

The bottom panel of Figure 2 shows logged traffic fatalities per capita, $\ln(K_i/L_i)$; cf. (20). Here I have assumed $\gamma < 1 + \beta$ to generate a negative relationship with population for $L_i < \hat{L}_i$, i.e., when all agents drive ($z_i = 1$). However, for $L_i > \hat{L}_i$ (meaning $z_i < 1$) the relationship is negative as long as

$$\gamma < \frac{1+\delta}{\delta - \lambda} + \beta, \tag{21}$$

which is a weaker condition than $\gamma < 1 + \beta$.

The main insight from Figure 2 is that the model can generate a pattern where more densely populated locations have more drivers, i.e., more congestion, but also lower rates of traffic fatalities, and a lower fraction of the population who drives.

Varying Income Levels Figure 3 shows the relationship between the fraction driving and population (i.e., z_i and L_i) for different income levels, y_i. The parameter values are otherwise the same as in the numerical example in Figure 2 (which refers to $y_i = 10$). The long-dashed black curve in Figure 3 thus corresponds to the curve in the second panel from the top in Figure 2.

For all levels of y_i, there is a negative relationship between the z_i and L_i. However, the relationship between incomes and driving is ambiguous. That is, locations with higher incomes may have higher, or lower, levels of z_i. For low population levels, the high-income location ($y_i = 10$, the long-dashed curve) drives the least (has the lowest z_i), but for higher population levels the lowest-income location ($y_i = 2.5$, the solid curve) has the lowest fraction driving.

Intuitively, there are two opposing effects involved. On the one hand, low incomes make the amount of public transit (P_i) smaller at any given tax rate, due to the smaller tax base, which makes provision of public transit less attractive. On the other hand, low incomes imply a high fixed cost of driving (D) relative to income, which makes voters more inclined to support public transit as an alternative to driving. Which effects dominates is ambiguous.

The same patterns as in Figure 3 hold for traffic fatality rates, since those vary (primarily) with z_i. In other words, the relationship between population density and traffic fatality rates (or the fraction driving) is unambiguously negative, while that between per-capita incomes and traffic fatalities (or the fraction driving) can be either positive or negative.

4.4 Migration

The analysis so far has treated L_i as exogenous. An alternative approach is to let agents migrate, so that L_i adjusts to equalize utilities across locations. Section A.2 in the Online Appendix explores a setting with two locations and frictionless migration (for simplicity assuming away agent heterogeneity, $\delta = 0$). Then there may exist multiple equilibria. Intuitively, since public transit is a public good, one location will tend to "specialize" in only driving, and the other provide some public transit; the location with larger population is the one with less driving. This aligns well with the findings when treating L_i as exogenous (and with the empirical evidence below), i.e., a negative relationship between population density and the fraction driving (and traffic deaths).

More interestingly, when allowing for differences in income between the two locations, then the location with the lower fraction driving could be either the richer, or the poorer, location. In that sense, the relationship between z_i and y_i is ambiguous also when allowing for migration, just as when we treated L_i as exogenous.

5 Data

The main empirical analysis makes use of two datasets, one made up by U.S. counties and the other by European regions. More discussion about sources and variable definitions are provided in Sections 5.1 and 5.2 below, with details deferred to Sections A.3.2 and A.3.3 of the Online Appendix.

As far as possible, all variables have been adjusted to be comparable between the two datasets. For example, population density is measured in persons per square kilometers and traffic deaths per 100,000 population. However, the units of measurement do not affect coefficient estimates, since all variables are logged in the empirical analysis.

The main difference between the two data sets is the year of study. For the U.S., the variables mostly refer to 2010, which is the latest year with good county-level population data. For the European data, I choose the year 2018, which is the latest with data available for the United Kingdom. The exception is European homicide rates, where I use the latest year available, which happens to be 2010.

I also use per-capita incomes for U.S. counties but GDP per capita for Europe, for reasons discussed in more detail in Section 5.2 below.

5.1 United States

U.S. data come from a few different sources. Population density and urbanization come from the 2010 U.S. Census, which is the latest year for which county-level population data is available. Population density is total county population divided by land area in square kilometers. The urbanization rate is computed as urban population over total population.

Since both density and urbanization are available only for 2010 (at least according to what I have found), for most other variables I try to use measures for the same year, or adjacent years.

Data on traffic fatalities are from the Fatality and Injury Reporting System Tool (FIRST) at the National Highway Traffic Safety Administration (NHTSA). I calculate the traffic fatality rate as total traffic fatalities in 2010 per $100,000$ population in the same year.

Population density in 1900 is based on data from a website maintained by Andrew J. Van Leuven (Van Leuven, 2020). This variable is also measured as people per square kilometer.

Per-capita income data are from the Bureau of Economic Analysis and refer to 2010.

Data on life expectancy are from the Centers for Disease Control and Prevention. These are measured in years and at birth and refer to 2010-2015.

Homicide (murder) rates are also computed using data from the Centers for Disease Control and Prevention, and measured as rates per 100,000 people. Since data is often missing I took the average rate over the decade 2006-2016 (2016 being the last year reported), allowing for better coverage than if using only a single year.

Data on the fraction commuters using public transit in 2010 come from the American Community Survey, conducted by the U.S. Census Bureau, and is calculated as the fraction using public transit out of the total number of commuters.

Data on the Republican vote share 2000-2016 and property taxes in 2010 come from Bazzi et al. (2020).

5.2 Europe

All data for Europe, and neighboring countries, come from an online database hosted by Eurostat, a body that compiles data from statistical agencies in EU member countries, and a few others (e.g., Turkey and Norway). I refer to these as "European" data for short.

I consider the regional disaggregation unit known as NUTS 2, which is the intermediate level of the Nomenclature of Territorial Units for Statistics. NUTS 2 regions are larger than most U.S. counties, which makes the European sample smaller, but also allows access to most variables needed. For example, traffic fatalities are not available at the more disaggregate NUTS 3 level.

As mentioned already, I use GDP per capita for Europe but income per capita for the U.S. The reason, in short, is that I have not been able to find per-capita income data from Eurostat. This is not ideal, as some inhabitants may work and produce in one region but to different degrees consume, drive, pay taxes, and/or use public transit in another region. However, this distinction may matter less for NUTS 2 regions than the much smaller U.S. counties, if GDP and income levels are more proportional to each other for larger economies than small.[7]

6 Empirical Analysis

6.1 Main Results

6.1.1 United States

Figure 4 plots log traffic fatality rates against log population density across all U.S. counties with data available, with state acronyms indicated. The negative relationship is clear. Counties in, say, New York are more densely populated and have lower traffic fatality rates than counties in, e.g., Texas and Nevada. However, the negative relationship is not driven (only) by differences between states or regions, but also appear in the same plots for the most populous states of California, Florida, New York, and Texas; see Figure 5.

Table 1 presents some regression results based on U.S. counties. The negative correlation is highly significant in the unconditional specification in column (1), corresponding to the plot in Figure 4. This holds when we add controls for log per-capita incomes and log life expectancy in columns (2) and (3); when we add state fixed effects in column (4); and when clustering standard errors on states in column (5).

[7]There are per-capita GDP data available for U.S. counties from the BEA, but for such small and open geographical units it arguably makes more sense to use per-capita incomes.

Controlling for life expectancy in columns (3)-(5) may be interpreted as partly absorbing variation in the quality of hospital care and related factors that decide if a vehicle collision is fatal, or not. The negative and significant coefficient on population density when controlling for life expectancy then suggests that this conditional correlation may be due to factors that cause (serious) crashes in the first place, and thus better capture the degree to which people rely and depend on cars.

Notably, controlling for life expectancy renders the coefficient on log per-capita incomes insignificant in Table 1, and it stays insignificant when adding state fixed effects. (The same holds when adding only state fixed effects, and no life expectancy control; those results are not reported here.) In other words, incomes are not as robustly associated with traffic deaths as density is, broadly consistent with the model's predictions.

Table A.1 in the Online Appendix presents the same regressions as in Table 1, but with urbanization in place of of population density. Those results are similar to the ones in Table 1, although the coefficient on log per-capita income there stays negative and significant.

6.1.2 Europe

Consider next European NUTS 2 regions. Figure 6 plots the log traffic fatality rate against log population density across all regions with data available, with country acronyms indicated, showing a clear negative relationship. In Figure 7, we also find similar negative patterns within France, Germany, Spain, and the UK, the countries with the largest numbers of NUTS 2 regions

The regression results in Table 2 show that the negative association between traffic fatalities and density is robust to similar controls as those that we used for the U.S. in Table 1. (As discussed, income per capita is here replaced by GDP per capita, due to data availability.) The other variables also show similar correlations as those found in U.S. data. Population density has a more robust association with traffic fatalities than GDP per capita (although it here takes country fixed effects to render the GDP per capita coefficient insignificant).

Table A.2 in the Online Appendix uses urbanization instead of population density, finding a negative association with traffic fatalities, again with more robust correlations for density than for GDP per capita.

6.1.3 Comparing the United States and Europe

As discussed in Section 5, not all variables are defined identically between the two data sets, and they also refer to different years: 2010 and 2018, for the U.S. and Europe, respectively. Note also that 2010 had comparatively low traffic fatality rates in the U.S. overall (see Section 3).

However, with these caveats in mind, it is interesting to note that the coefficients on the density variable are similar in size when comparing Tables 1 and 2, in particular when including fixed effects.

To illustrate this, Figure 8 shows a single plot of log traffic fatalities against log population density where U.S. counties and European regions are overlaid, together with associated best-fit lines. The intercepts differ, with the U.S. being more deadly, but the slopes are very similar.[8] The slopes become even more similar if we drop sparsely populated but relatively safe Norway. My interpretation is that these correlations reflect a relatively deep-rooted force that is common across otherwise quite different societies. Urban and densely populated places always tend to be less car dependent.

At the same time, the fact that the intercept is so much higher for the U.S. shows that there are differences in traffic fatality rates across the Atlantic. To gauge this quantitatively, I used the unconditional regression in column (1) of Table 1 to predict the U.S. traffic death rate based on the mean of the log population density in Europe (unweighted across all regions). This generates a drop in the mean (non-logged) U.S. traffic fatality rate from 23.1 to 9.6 per 100,000 people, which can be compared to a sample mean of 5.6 for Europe. In other words, based on this regression, a bit over half of the U.S.-Europe gap in traffic fatality rates can be accounted for by differences in population density.

[8] I do not show the results here, but if we run regressions on these merged data, then the coefficient on a U.S. dummy always comes out as positive and significant, while the interaction between U.S. dummy and log population density tends to either come out as insignificant, or, when more precisely estimated, small in size relative to the coefficient on log population density.

6.2 Other Measures of Car Dependence

This paper interprets traffic fatalities as a proxy for (or at least something connected to) car dependence in a broad sense. Next, I consider a couple of alternative measures. I have not been able to find a single measure that is available both across U.S. counties and European regions. For the U.S., I consider the fraction of the commuting population who relied on public transit, and for Europe the number of vehicles per capita.

The latter of these might be most in line with the existing literature.[9] Alas, while vehicles per capita can be found at the state level in the U.S. (and for select cities) it is apparently not available at the county level.

Table 3 presents results for the U.S., based on the exact same type of regressions as in Table 1, but with the log transit use rate as the dependent variable. The correlations now carry the opposite sign, as one would expect, since more public transit use is associated less car dependence. The estimated coefficient on log population density comes out as positive and highly significant throughout, while the results for per-capita incomes are slightly weaker, consistent with the results in Table 1.

There are also some notable differences. The sample size and the explanatory power are both much higher when using traffic fatalities in Table 1 than with public transit use in Table 3; the sample shrinks by about 250 counties and R-squared falls from the .41-.48 range to .10-.31. The estimated coefficients are also somewhat smaller in absolute terms. Although one can have different interpretations, my own suggested conclusion is that the traffic death rate better captures variation in car dependence than public transit use does, at least in a U.S. context.

Table 4 presents the same type of regressions across European NUTS 2 regions as those in Table 2, but with log vehicles per capita as the dependent variable, instead of the log traffic fatality rate. In column (1), we note that high population density is associated with more vehicles per capita, the opposite of what we would expect given our previous result. However, this is driven by urban areas being richer. When controlling for per-capita GDP the correlations come out as negative and significant, which holds also with the other

[9]See, e.g., Ostermeijer et al. (2022), who use vehicles per capita at the city level worldwide.

controls, and qualitatively matches the results for traffic fatality rates in Table 2. However, the coefficient on log GDP per capita comes out as positive, the exact opposite of what we found in Table 2. Again, this conforms with the previous findings, and with the model, that income (or GDP) has a less consistent association with car dependence than population density does.

Moreover, like with the transit use rate in the U.S. data, the sample here is smaller due to vehicle data being more limited, and the R-squared statistic mostly comes out as lower. The magnitude of the estimated elasticities differs by a factor of almost 4. Considering the specification in column (4) of each table, a 1% increase in population density is associated with roughly a .31% decline in traffic fatality rates in Table 2 but only a .083% drop in vehicles per capita in Table 4.

6.3 Mechanisms

6.3.1 Homicide Rates

One possibility is that high rates of traffic fatalities are caused by rural areas having a more violent (and/or careless) culture, and thus less safe driving. For example, Grosjean (2014) and Couttenier et al. (2017) link violence in the U.S. to deep-rooted cultural and institutional events, which could affect traffic behavior too.

To explore this possibility, I next control for homicide rates, a common proxy for overall violence (see, e.g., Pinker, 2011). This shrinks the samples, due to missing homicide data for many counties and NUTS 2 regions.[10] However, the two data sets still cover some counties/regions in at least 44 U.S. states and 21 European countries, respectively.

Table 5 presents the same regressions as in Table 1, but adding a control for the log homicide rate. The coefficient on log population density stays negative and highly significant, and the coefficient on the log homicide rate comes out as positive and almost equally significant throughout. This suggests that what is often called traffic violence may be associated with more general forms of violence.

Table 6 considers Europe, adding a control for the log homicide rate to the regressions

[10] For Europe we also lose four observations with no homicides when logging.

in Table 2. Like for the U.S., the negative association between population density and traffic deaths is unchanged when adding this control, and the homicide rate itself shows a positive correlation with traffic deaths, although here the estimates are not significant when adding country fixed effects.

In the Online Appendix, I present results when using the log homicide rate as the dependent variable (see Tables A.3 and A.4). For the U.S., homicide rates are indeed lower in more densely populated counties, although the estimated coefficient estimates are not always significant. For Europe, the association between homicide rates and density tends to come out as *positive*, so rural parts of Europe are not generally more violent than cities, even though they have worse road safety.

In sum, there is little to suggest that the link between population density and traffic fatalities captures a culture of violence. However, the homicide rate itself tends to show a positive association with traffic fatalities (in particular in U.S. data) so there seems to exist some connection between traffic violence and other forms of violence.

6.3.2 Historical Population Density

Next I explore if the results could be driven by factors affecting both traffic fatalities and city growth and thus population density. For example, car dependence may lead to both traffic fatalities and urban sprawl, and thus low density. Ostermeijer et al. (2022) compare cities across the world and find that car ownership, instrumented by the presence of domestic car manufacturing in 1920, appears to lead to lower urban population density. U.S. highways also seem to have had a causal effect on urban sprawl (Baum-Snow, 2007, 2010).

To explore to what extent this chain of causation may drive our results, Table 7 considers the sample of U.S. counties but regresses the log traffic fatality rate on log population density in 1900 instead of modern population density, prior to the expansion of car use. (I have not been able to find any historical population density for European NUTS 2 regions.) The correlations are still negative and highly significant in all specifications, although slightly smaller in size compared to Table 1. Interestingly, the coefficient on per-capita incomes in Table 7 comes out as more consistently significant, possibly because it helps absorb some of the variation captured by modern population density in Table 1.

Table A.5 in the Online Appendix instead uses population density in 1900 as instrument for modern population density. The coefficient on the instrumented density variable is negative and highly significant throughout, while that on per-capita incomes turns insignificant when adding state fixed effects.

Absent other factors affecting both population density in 1900 and modern traffic fatality rates, this might suggest that the link is causal.

6.3.3 Political Covariates

Table 8 presents results from the same U.S. county level-regressions as in Table 1, but controlling for two political variables: the Republican Presidential vote share 2000-2016 and the property tax rate in 2010, both logged. These may absorb some of the mechanisms through which population density can impact traffic fatalities. For example, higher density may make voters more supportive of property taxes to fund public transit, making people less car dependent and traffic fatalities lower, and the Republican vote share may proxy for a general resistance to such taxes.

However, entering these controls does not render the coefficient on log population density insignificant, although it shrinks the estimated size of the coefficient marginally. (This also holds when letting each of the two controls enter separately.) This is actually consistent with the model, where traffic deaths (and the fraction driving) are decreasing with population density also over those intervals where tax rates stay constant; cf. Figure 2.

Moreover, both of these two political variables seem to carry the expected signs when significant: traffic fatality rates tend to be higher in counties with a more right-leaning voting population, and lower property tax rates. If we think of these variables as capturing a general willingness to fund public transit, this seems consistent with the model. However, neither variable comes out as significant when entering state fixed effects.

6.3.4 Kitchen-Sink Regressions

The final exercise I undertake is to regress traffic death rates on population density, per-capita incomes (or GDP), and life expectancy, as well as the other variables considered in the previous sections. For the U.S., these are the homicide rate, the fraction using public transit, the Republican Presidential vote share, and the property tax rate (all logged as before). For Europe, I use the homicide rate, vehicles per capita, and one more variable that I have not explored yet, the length in kilometres of motorways per capita (all logged).

This kind of "kitchen-sink" approach allows us to explore if anything used thus far can render population density insignificant.

Results from a few such regressions are presented in Tables 9 and 10, for the U.S. and Europe, respectively. In both tables, I restrict the sample to those counties and NUTS 2 regions for which we have data on all variables, meaning all regressions are based on the same minimum sample (425 counties and 119 NUTS 2 regions). All regressions include state fixed effects for the U.S. and country fixed effects for Europe, and standard errors are clustered on states and countries, respectively. These are arguably the most "demanding" specifications we may consider using these data.

The results in both Tables 9 and 10 show that the only variable that stays significant throughout at the 1% level is log population density.

As in Section 6.3.1, the log homicide rate carries a positive coefficient and is relatively precisely estimated in the U.S. data, although not for Europe.

The property tax rate comes out with the expected negative sign for the U.S. in Table 9, here indeed a little stronger than in Table 8. As discussed already, this matches the model well.

For Europe, log vehicles per capita comes out as positive and significant at the 10% level, which is hardly surprising. But, as mentioned, this does not weaken the estimated coefficient on log population density, so the link between density and road safety seems to be about more than just car ownership.

These results are qualitatively very similar when not restricting the samples, in particular the significance levels for the negative coefficient estimates for log population density.

7 Conclusion

This paper has explored the empirical relationship between population density and traffic death rates, finding a very robust negative correlation across both U.S. counties and European regions. Urban areas have better traffic safety records than rural areas in both Europe and the U.S. This contrasts with other externalities from driving, such as congestion and air pollution, which are well known to be worse in cities.

While the results for population density are very robust, the relationship between per-capita incomes and traffic death rates is more mixed, although mostly negative.

To make sense of some of these patterns I have also presented a model where agents choose between driving and using tax-funded public transit. Once a location becomes sufficiently densely populated, it switches from a zero-tax and zero-public-transit mode to a regime with some taxation and public transit use, although drivers are still in majority. Further increases in density can eventually make non-drivers a majority, rasing tax rates and public transit provision further.

The model can explain why higher population density is associated with a larger total number of drivers (i.e., more congestion and pollution), but a lower fraction of the population driving, and lower traffic fatality rates.

At the same time, the effects of changes to per-capita incomes are ambiguous, consistent with the somewhat weaker correlation between traffic fatality rates and per-capita incomes found in the data.

Acknowledgements

This research project was supported in part by the Social Sciences and Humanities Research Council of Canada.

References

Akbar, P., Couture, V., Duranton, G., and Storeygard, A. (2023). Mobility and Congestion in Urban India. *American Economic Review*, 113(4):1083–1111.

Albalate, D. and Fageda, X. (2021). On the relationship between congestion and road safety in cities. *Transport Policy*, 105:145–152.

Anderson, M. (2008). Safety for whom? The effects of light trucks on traffic fatalities. *Journal of Health Economics*, 27(4):973–989.

Anderson, M. L. (2014). Subways, strikes, and slowdowns: the impacts of public transit on traffic congestion. *American Economic Review*, 104(9):2763–96.

Anderson, M. L. and Auffhammer, M. (2014). Pounds That Kill: The External Costs of Vehicle Weight. *The Review of Economic Studies*, 81(2):535–571.

Basso, L. J., Navarro, M., and Silva, H. E. (2021). Public transport and urban structure. *Economics of Transportation*, 28:100232.

Bauernschuster, S. and Rekers, R. (2022). Speed limit enforcement and road safety. *Journal of Public Economics*, 210:104663.

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

Baum-Snow, N. (2010). Changes in transportation infrastructure and commuting patterns in US metropolitan areas, 1960-2000. *American Economic Review*, 100(2):378–82.

Bazzi, S., Fiszbein, M., and Gebresilasse, M. (2020). Frontier culture: the roots and persistence of "rugged individualism" in the United States. *Econometrica*, 88(6):2329–2368.

Beland, L. P. and Brent, D. A. (2018). Traffic and crime. *Journal of Public Economics*, 160:96–116.

Chalfin, A. and Massenkoff, M. N. (2022). A New Racial Disparity in Traffic Fatalities. *NBER Working Paper*, 30636.

Couttenier, M., Grosjean, P., and Sangnier, M. (2017). The Wild West IS Wild: The Homicide Resource Curse. *Journal of the European Economic Association*, 15(3):558–585.

De Borger, B. and Proost, S. (2012). A political economy model of road pricing. *Journal of Urban Economics*, 71(1):79–92.

De Borger, B. and Proost, S. (2015). The political economy of public transport pricing and supply decisions. *Economics of Transportation*, 4(1-2):95–109.

Duranton, G. and Turner, M. A. (2018). Urban form and driving: Evidence from US cities. *Journal of Urban Economics*, 108:170–191.

Duranton, G., Turner, M. A., Arnott, R., Borenstein, S., Desmet, K., Brueckner, J., Couture, V., Glaeser, E., Kohlhagen, S., Kopp, A., Levinson, D., Michaels, G., and Mun, K. (2011). The Fundamental Law of Road Congestion: Evidence from US Cities. *American Economic Review*, 101(6):2616–52.

French, M. T. and Gumus, G. (2021). Death on the job: The Great Recession and work-related traffic fatalities. *Social Science and Medicine*, 280:113979.

Green, C. P., Heywood, J. S., and Navarro, M. (2016). Traffic accidents and the London congestion charge. *Journal of Public Economics*, 133:11–22.

Grosjean, P. (2014). A history of violence: the culture of honor and homicide in the U.S. South. *Journal of the European Economic Association*, 12(5):1285–1316.

Hughes, J. E., Kaffine, D., and Kaffine, L. (2023). Decline in traffic congestion increased crash severity in the wake of COVID-19. *Transportation Research Record*, 2677(4):892–903.

Jacobsen, M. R. (2013). Fuel economy and safety: the influences of vehicle class and driver behavior. *American Economic Journal: Applied Economics*, 5(3):1–26.

Kenworthy, J. R. and Laube, F. B. (1999). Patterns of automobile dependence in cities: an international overview of key physical and economic dimensions with some implications for urban policy. *Transportation Research Part A: Policy and Practice*, 33(7-8):691–723.

Maheshri, V. and Winston, C. (2016). Did the Great Recession keep bad drivers off the road? *Journal of Risk and Uncertainty*, 52(3):255–280.

McIntosh, J., Trubka, R., Kenworthy, J., and Newman, P. (2014). The role of urban form and transit in city car dependence: Analysis of 26 global cities from 1960 to 2000. *Transportation Research Part D: Transport and Environment*, 33:95–110.

Newman, P. W. and Kenworthy, J. R. (1989). Gasoline consumption and cities. *Journal of the American Planning Association*, 55(1):24–37.

Ostermeijer, F., Koster, H. R. A., van Ommeren, J., and Nielsen, V. M. (2022). Automobiles and urban density. *Journal of Economic Geography*, 22(5):1073–1095.

Parry, I. W. and Small, K. A. (2005). Does Britain or the United States Have the Right Gasoline Tax? *American Economic Review*, 95(4):1276–1289.

Parry, I. W. and Small, K. A. (2009). Should Urban Transit Subsidies Be Reduced? *American Economic Review*, 99(3):700–724.

Pinker, S. (2011). *The Better Angels of our Nature: Why Violence Has Declined*. Viking Press, New York.

Retallack, A. E. and Ostendorf, B. (2019). Current understanding of the effects of congestion on traffic accidents. *International Journal of Environmental Research and Public Health*, 16(18):3400.

Ruhm, C. J. (1995). Economic conditions and alcohol problems. *Journal of Health Economics*, 14(5):583–603.

Ruhm, C. J. (2000). Are Recessions Good for Your Health? *The Quarterly Journal of Economics*, 115(2):617–650.

Schönholzer, D., Lane, G., and Kelley, E. (2022). Information and strategy in lemon markets: Improving safety in informal transit. *mimeo, IIES, Stockholm University*.

Seidel, T. and Wickerath, J. (2020). Rush hours and urbanization. *Regional Science and Urban Economics*, 85:103580.

Shefer, D. and Rietveld, P. (1997). Congestion and safety on highways: Towards an analytical model. *Urban Studies*, 34(4):679–692.

Tang, C. K. and Van Ommeren, J. (2022). Accident externality of driving: evidence from the London Congestion Charge. *Journal of Economic Geography*, 22(3):547–580.

Tyndall, J. (2021). Pedestrian deaths and large vehicles. *Economics of Transportation*, 26-27:100–219.

Van Leuven, A. J. (2020). Historical U.S. County Population Data: 1900 to 2010 https://andrewvanleuven.com/post/countypopulations.

Vickrey, W. (1980). Optimal transit subsidy policy. *Transportation*, 9(4):389–409.

Wright, N. A. and Dorilas, E. (2022). Do Cellphone Bans Save Lives? Evidence From Handheld Laws on Traffic Fatalities. *Journal of Health Economics*, 85:102659.

Yasin, Y. J., Grivna, M., and Abu-Zidan, F. M. (2021). Global impact of COVID-19 pandemic on road traffic collisions. *World Journal of Emergency Surgery*, 16(1):1–14.

Tables and Figures

	Dependent variable is the log traffic fatality rate				
	(1)	(2)	(3)	(4)	(5)
Log population density	-0.304^{***}	-0.295^{***}	-0.307^{***}	-0.343^{***}	-0.343^{***}
	(0.007)	(0.008)	(0.008)	(0.010)	(0.014)
Log income per capita		-0.425^{***}	-0.060	0.068	0.068
		(0.061)	(0.067)	(0.078)	(0.104)
Log life expectancy			-4.623^{***}	-1.831^{***}	-1.831^{***}
			(0.438)	(0.540)	(0.631)
R^2	0.41	0.42	0.43	0.48	0.48
Number of obs.	2866	2810	2717	2717	2717
Fixed effects	None	None	None	State	State
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. Robust standard errors are indicated in parentheses, except for column (5), which clusters on state. The unit of observation is a county. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 1: Traffic Deaths and Population Density: United States.

	Dependent variable is the log traffic fatality rate				
	(1)	(2)	(3)	(4)	(5)
Log population density	−0.252***	−0.149***	−0.159***	−0.310***	−0.310***
	(0.026)	(0.030)	(0.031)	(0.033)	(0.051)
Log GDP per capita		−0.555***	−0.459***	0.112	0.112
		(0.082)	(0.091)	(0.092)	(0.135)
Log life expectancy			−2.440***	−1.944	−1.944
			(0.696)	(1.851)	(2.491)
R^2	0.34	0.48	0.50	0.80	0.80
Number of obs.	317	274	271	271	271
Fixed effects	None	None	None	Country	Country
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. Robust standard errors are indicated in parentheses, except for column (5), which clusters on country. The unit of observation is a NUTS 2 region. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 2: Traffic Deaths and Population Density: Europe.

	Dependent variable is the log public transit use rate				
	(1)	(2)	(3)	(4)	(5)
Log population density	0.237***	0.197***	0.242***	0.266***	0.266***
	(0.019)	(0.019)	(0.020)	(0.022)	(0.046)
Log income per capita		1.278***	0.569***	0.323*	0.323
		(0.131)	(0.160)	(0.167)	(0.225)
Log life expectancy			7.449***	1.996*	1.996
			(0.908)	(1.089)	(1.480)
R^2	0.10	0.14	0.17	0.31	0.31
Number of obs.	2581	2432	2355	2355	2355
Fixed effects	None	None	None	State	State
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. The dependent variable is the log of the fraction using public transit out of those that either drive or use public transit. (Driving includes car pooling.) Robust standard errors are indicated in parentheses, except for column (5), which clusters on state. The unit of observation is a county. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 3: Public Transit Rates and Population Density: United States.

	Dependent variable is log vehicles per capita				
	(1)	(2)	(3)	(4)	(5)
Log population density	0.034*	−0.071***	−0.058**	−0.083***	−0.083**
	(0.019)	(0.024)	(0.023)	(0.024)	(0.037)
Log GDP per capita		0.776***	0.607***	0.535***	0.535**
		(0.099)	(0.114)	(0.120)	(0.216)
Log life expectancy			2.899***	−1.838	−1.838
			(0.703)	(1.708)	(1.830)
R^2	0.01	0.42	0.44	0.83	0.83
Number of obs.	293	251	249	249	249
Fixed effects	None	None	None	Country	Country
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. The dependent variable is the log of vehicles per capita. Robust standard errors are indicated in parentheses, except for column (5), which clusters on state. The unit of observation is a NUTS 2 region. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4: Vehicles per Capita and Population Density: Europe.

	Dependent variable is the log traffic fatality rate				
	(1)	(2)	(3)	(4)	(5)
Log population density	−0.246***	−0.218***	−0.205***	−0.208***	−0.208***
	(0.015)	(0.017)	(0.015)	(0.016)	(0.020)
Log income per capita		−0.361***	−0.150	−0.100	−0.100
		(0.114)	(0.108)	(0.109)	(0.173)
Log life expectancy			−4.464***	−4.466***	−4.466***
			(1.025)	(1.303)	(1.376)
Log homicide rate	0.279***	0.247***	0.157***	0.126***	0.126**
	(0.035)	(0.035)	(0.044)	(0.045)	(0.051)
R^2	0.56	0.58	0.59	0.72	0.72
Number of obs.	469	460	430	430	430
Fixed effects	None	None	None	State	State
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. Robust standard errors are indicated in parentheses, except for column (5), which clusters on state. The unit of observation is a county. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 5: Controlling for Homicide Rates: United States.

	Dependent variable is the log traffic fatality rate				
	(1)	(2)	(3)	(4)	(5)
Log population density	-0.192^{***}	-0.133^{***}	-0.134^{***}	-0.326^{***}	-0.326^{***}
	(0.035)	(0.032)	(0.032)	(0.044)	(0.052)
Log GDP per capita		-0.466^{***}	-0.369^{***}	0.160	0.160
		(0.104)	(0.116)	(0.129)	(0.194)
Log life expectancy			-2.883^{***}	-3.077	-3.077
			(0.940)	(2.349)	(2.393)
Log homicide rate	0.433^{***}	0.240^{***}	0.208^{***}	0.065	0.065
	(0.047)	(0.051)	(0.052)	(0.042)	(0.047)
R^2	0.43	0.54	0.58	0.85	0.85
Number of obs.	159	154	153	153	153
Fixed effects	None	None	None	Country	Country
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. Robust standard errors are indicated in parentheses, except for column (5), which clusters on country. The unit of observation is a NUTS 2 region. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 6: Controlling for Homicide Rates: Europe.

	Dependent variable is the log traffic fatality rate				
	(1)	(2)	(3)	(4)	(5)
Log population density 1900	-0.208^{***}	-0.196^{***}	-0.202^{***}	-0.229^{***}	-0.229^{***}
	(0.012)	(0.012)	(0.012)	(0.018)	(0.042)
Log income per capita		-0.965^{***}	-0.738^{***}	-0.798^{***}	-0.798^{***}
		(0.073)	(0.084)	(0.094)	(0.093)
Log life expectancy			-3.289^{***}	-0.944	-0.944
			(0.543)	(0.684)	(0.947)
R^2	0.16	0.23	0.24	0.31	0.31
Number of obs.	2602	2566	2493	2493	2493
Fixed effects	None	None	None	State	State
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. Robust standard errors are indicated in parentheses, except for column (5), which clusters on state. The unit of observation is a county. * indicates p <0.10, ** p <0.05, and *** p <0.01.

Table 7: Modern Traffic Deaths and Population Density in 1900: United States.

	Dependent variable is the log traffic fatality rate				
	(1)	(2)	(3)	(4)	(5)
Log population density	-0.280^{***}	-0.273^{***}	-0.290^{***}	-0.340^{***}	-0.340^{***}
	(0.008)	(0.008)	(0.009)	(0.011)	(0.016)
Log income per capita		-0.297^{***}	-0.039	0.062	0.062
		(0.062)	(0.068)	(0.079)	(0.105)
Log life expectancy			-4.094^{***}	-1.838^{***}	-1.838^{***}
			(0.466)	(0.543)	(0.644)
Log poperty tax rate	-0.200^{***}	-0.165^{***}	-0.094^{***}	-0.050	-0.050
	(0.022)	(0.023)	(0.025)	(0.056)	(0.082)
Log Republican vote share	0.195^{***}	0.198^{***}	0.152^{***}	-0.010	-0.010
	(0.051)	(0.051)	(0.051)	(0.059)	(0.079)
R^2	0.43	0.43	0.44	0.48	0.48
Number of obs.	2842	2793	2707	2707	2707
Fixed effects	None	None	None	State	State
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. Robust standard errors are indicated in parentheses, except for column (5), which clusters on state. The unit of observation is a county. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 8: Traffic Deaths and Political Variables: United States.

	Dependent variable is the log traffic fatality rate				
	(1)	(2)	(3)	(4)	(5)
Log population density	−0.210***	−0.205***	−0.202***	−0.205***	−0.175***
	(0.023)	(0.020)	(0.025)	(0.023)	(0.025)
Log income per capita	−0.091	−0.086	−0.117	−0.086	−0.114
	(0.174)	(0.174)	(0.161)	(0.173)	(0.146)
Log life expectancy	−6.444***	−4.758***	−6.417***	−6.431***	−4.299***
	(1.453)	(1.479)	(1.408)	(1.448)	(1.522)
Log homicide rate		0.111*			0.152**
		(0.056)			(0.069)
Log poperty tax rate			−0.146*		−0.135*
			(0.083)		(0.076)
Log Republican vote share			0.041		0.145
			(0.083)		(0.109)
Log public transit rate				−0.009	−0.007
				(0.018)	(0.023)
R^2	0.71	0.71	0.71	0.71	0.72
Number of obs.	425	425	425	425	425
Fixed effects	State	State	State	State	State
Standard errors	Clustered	Clustered	Clustered	Clustered	Clustered

Notes: Ordinary least squares regressions. Standard errors are indicated in parentheses, all clustered on state. All specifications also include state fixed effects. All columns are based on the minimum sample for which we have data on all variables. The unit of observation is a county. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 9: The Kitchen Sink: United States.

	Dependent variable is the log traffic fatality rate				
	(1)	(2)	(3)	(4)	(5)
Log population density	-0.363^{***}	-0.367^{***}	-0.329^{***}	-0.353^{***}	-0.325^{***}
	(0.033)	(0.032)	(0.028)	(0.032)	(0.025)
Log GDP per capita	0.148	0.151	-0.059	0.153	-0.041
	(0.130)	(0.142)	(0.183)	(0.114)	(0.189)
Log life expectancy	-2.864	-2.143	-2.792	-3.293^{*}	-2.303
	(1.939)	(1.955)	(1.974)	(1.720)	(1.929)
Log homicide rate		0.052			0.065
		(0.040)			(0.042)
Log vehicles per capita			0.424^{**}		0.408^{**}
			(0.177)		(0.188)
Log km motorway per capita				0.042	0.040
				(0.035)	(0.029)
R^2	0.86	0.86	0.87	0.86	0.87
Number of obs.	119	119	119	119	119
Fixed effects	Country	Country	Country	Country	Country
Standard errors	Clustered	Clustered	Clustered	Clustered	Clustered

Notes: Ordinary least squares regressions. Standard errors are indicated in parentheses, all clustered on country. All specifications also include country fixed effects. All columns are based on the minimum sample for which we have data on all variables. The unit of observation is a NUTS 2 region. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 10: The Kitchen Sink: Europe.

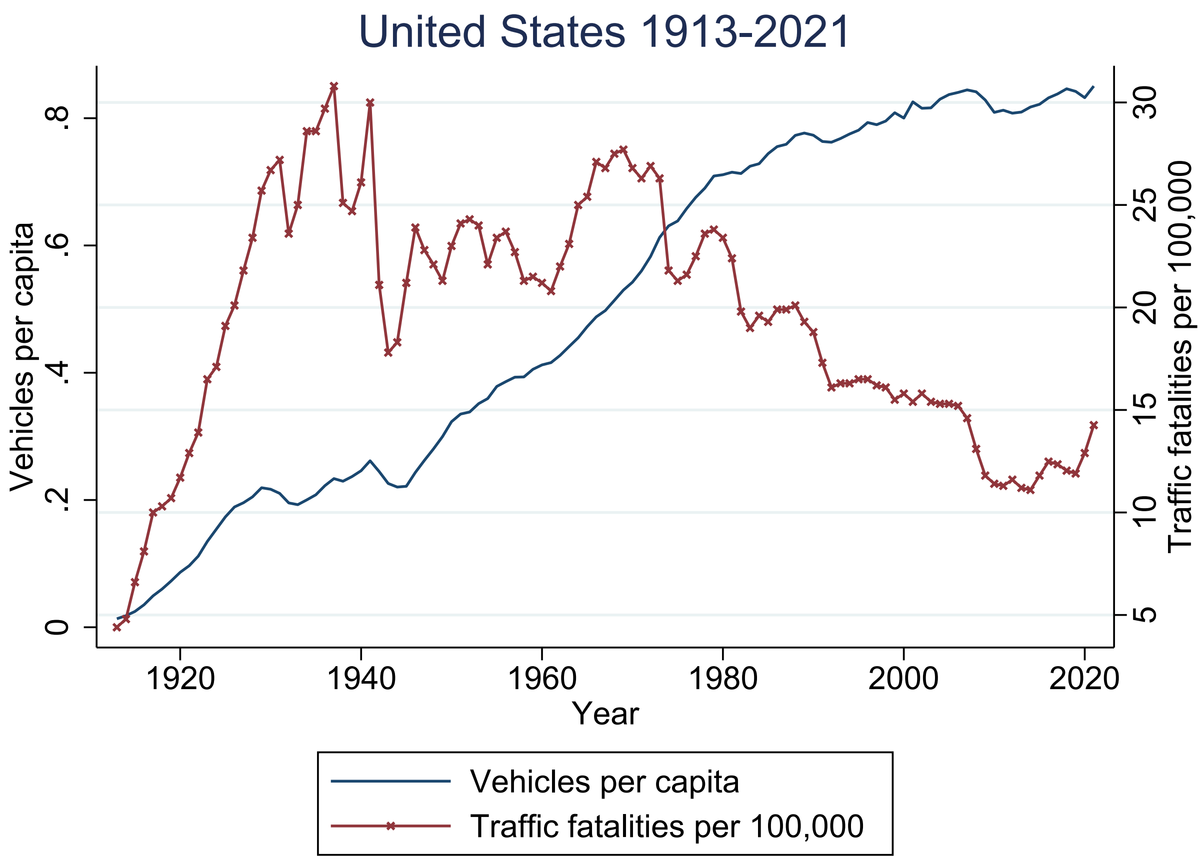

Figure 1: Aggregate time trends in the United States.

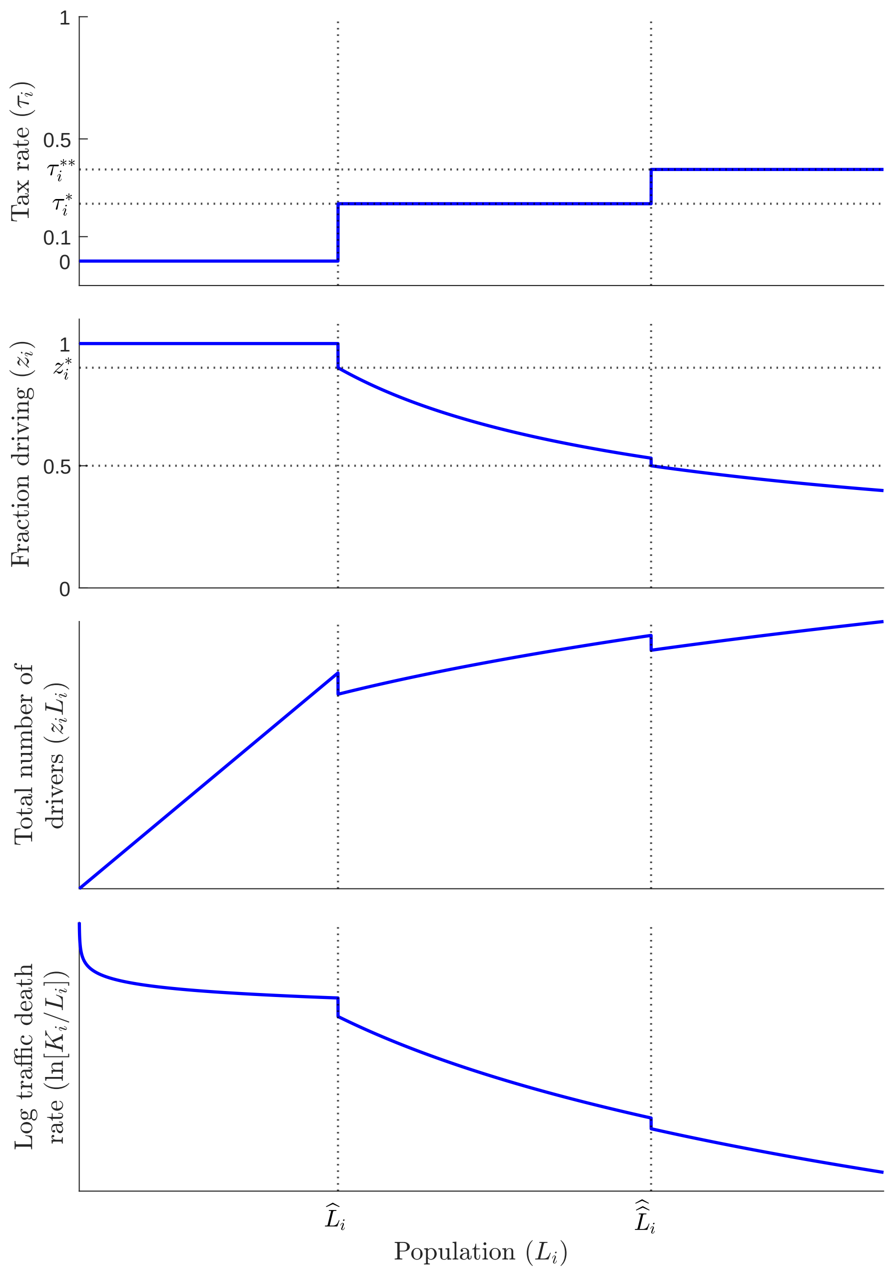

Figure 2: Illustration of the model.

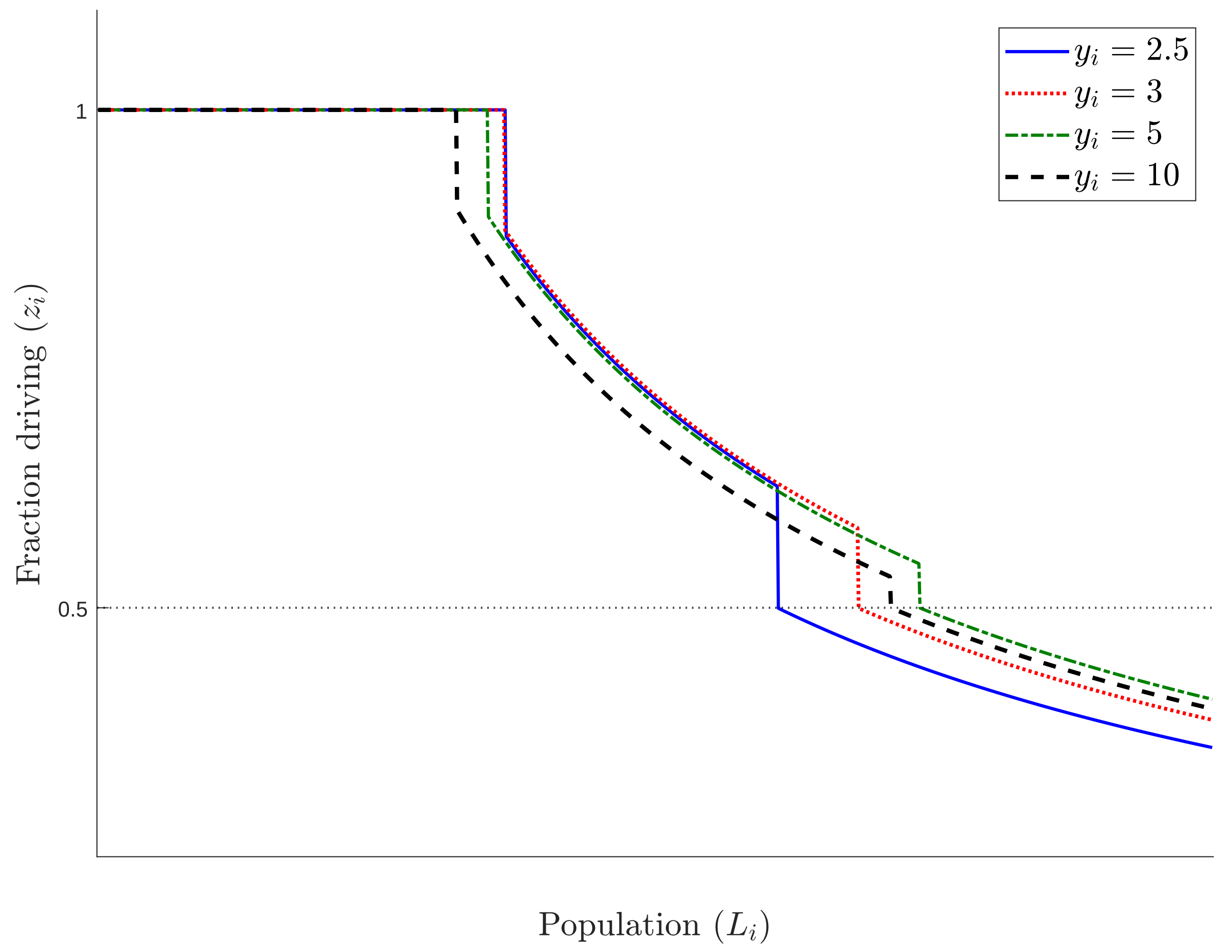

Figure 3: Varying income levels (y_i).

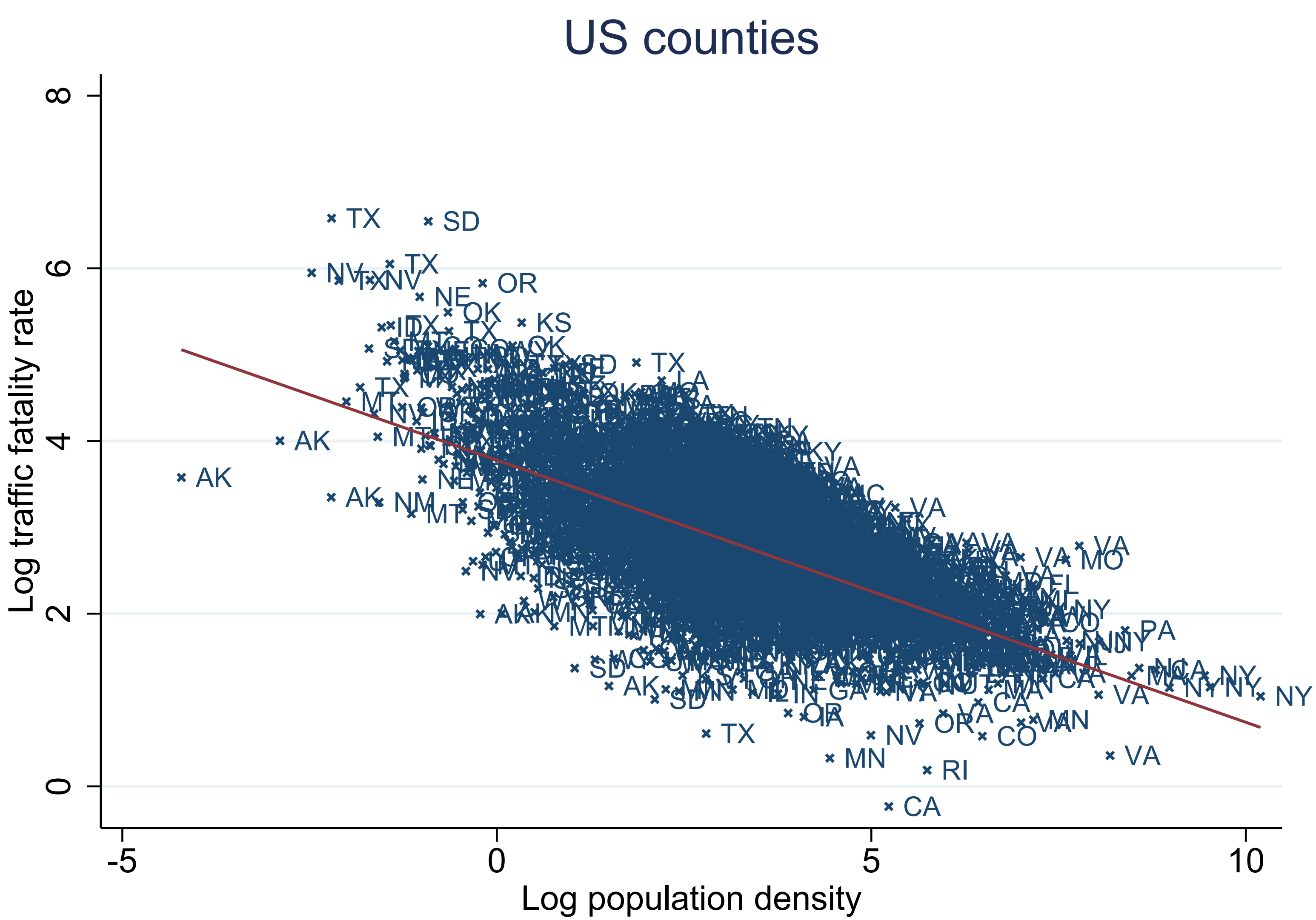

Figure 4: Traffic fatality rates and population densities across all U.S. counties.

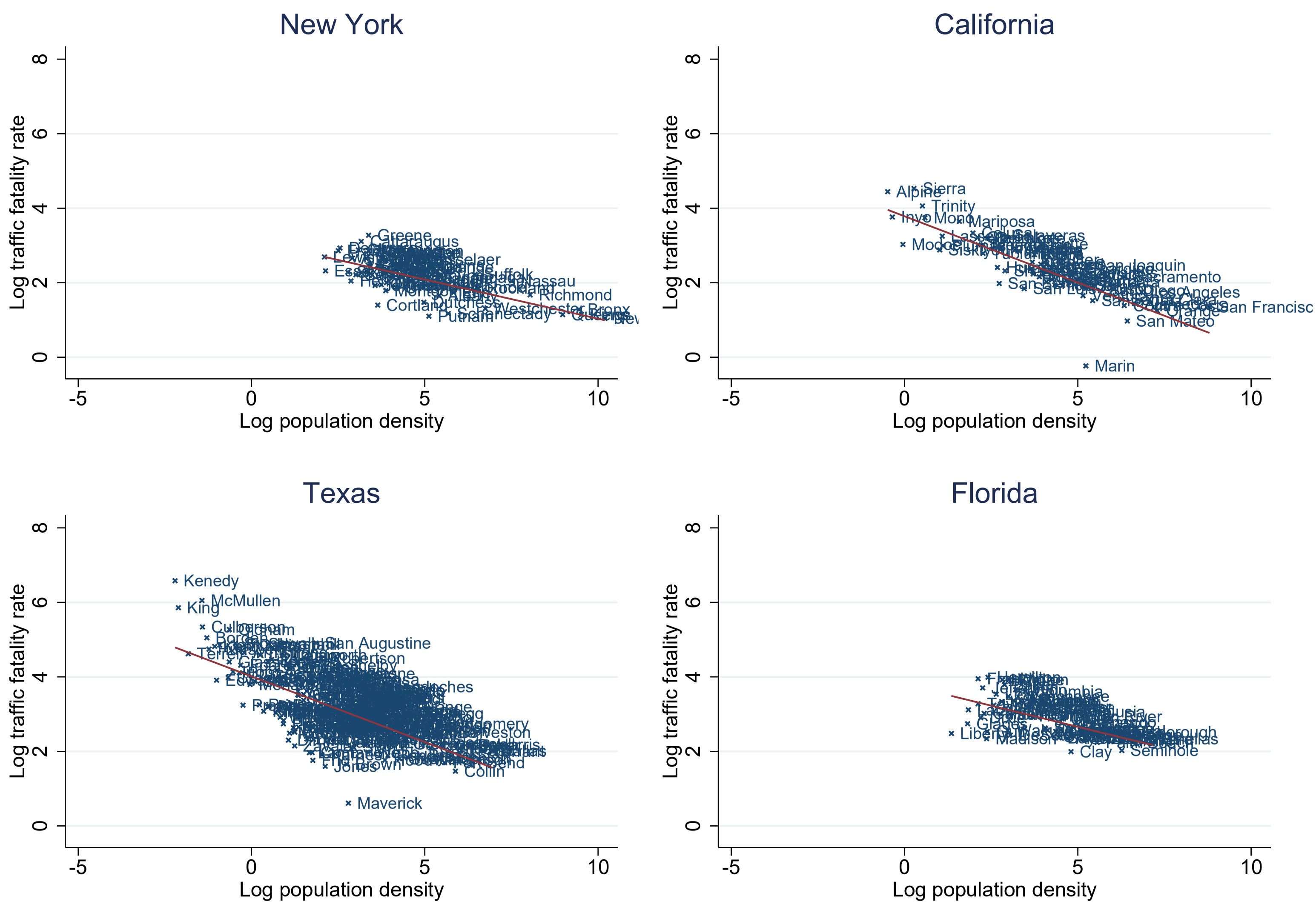

Figure 5: Traffic fatality rates and population densities across counties in four U.S. states.

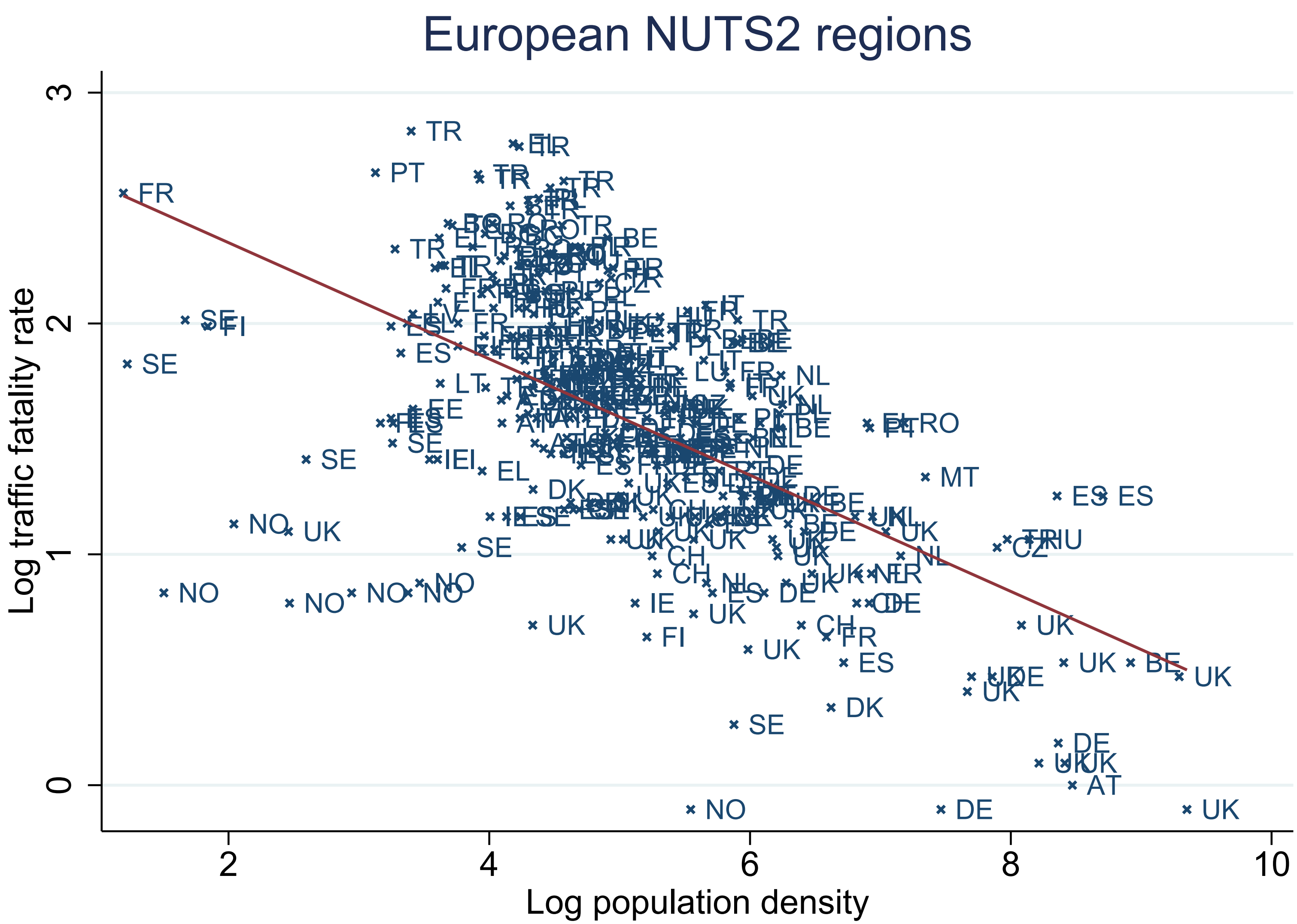

Figure 6: Traffic fatality rates and population densities across all NUTS 2 regions.

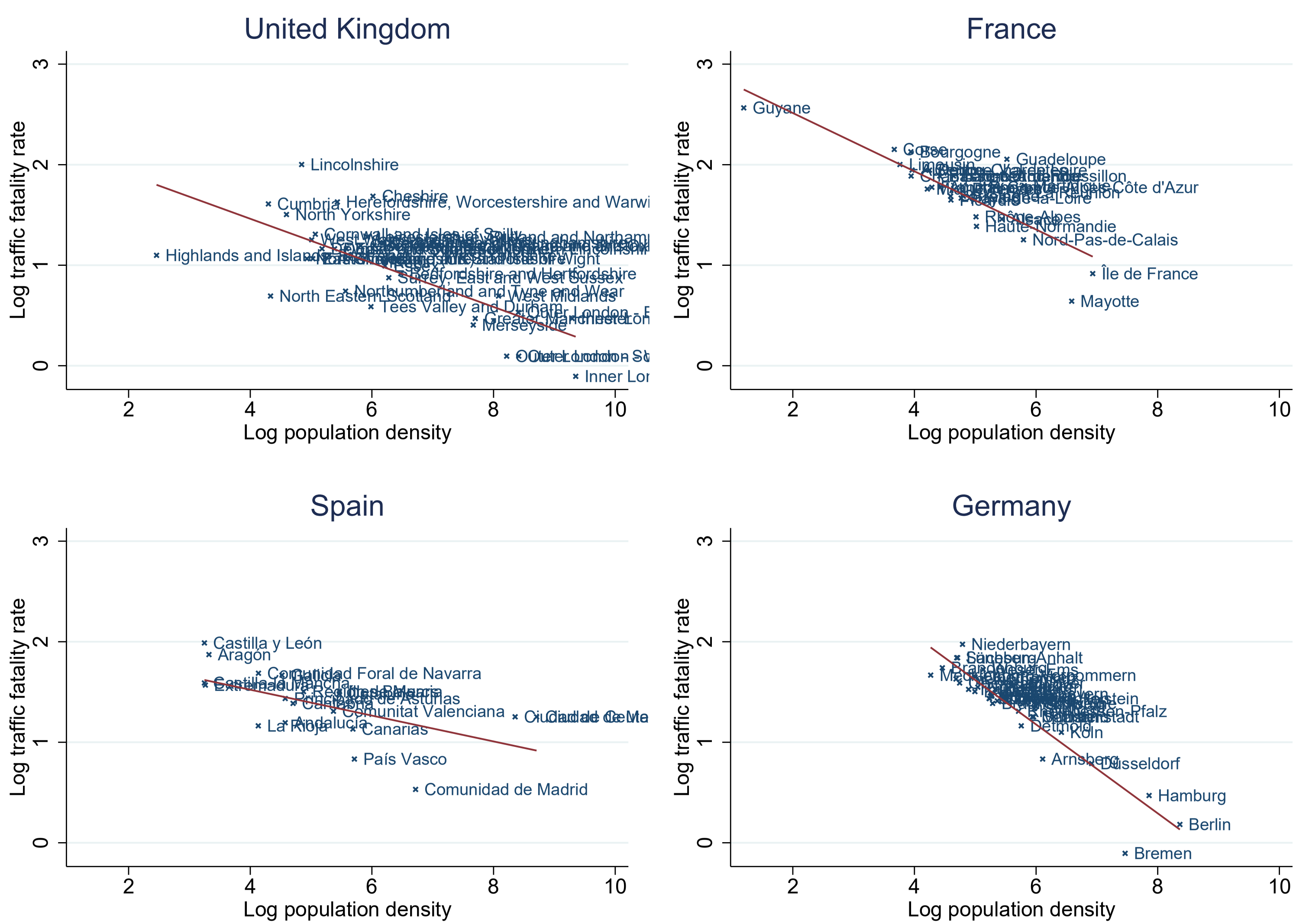

Figure 7: Traffic fatality rates and population densities across NUTS 2 regions in four Western European countries.

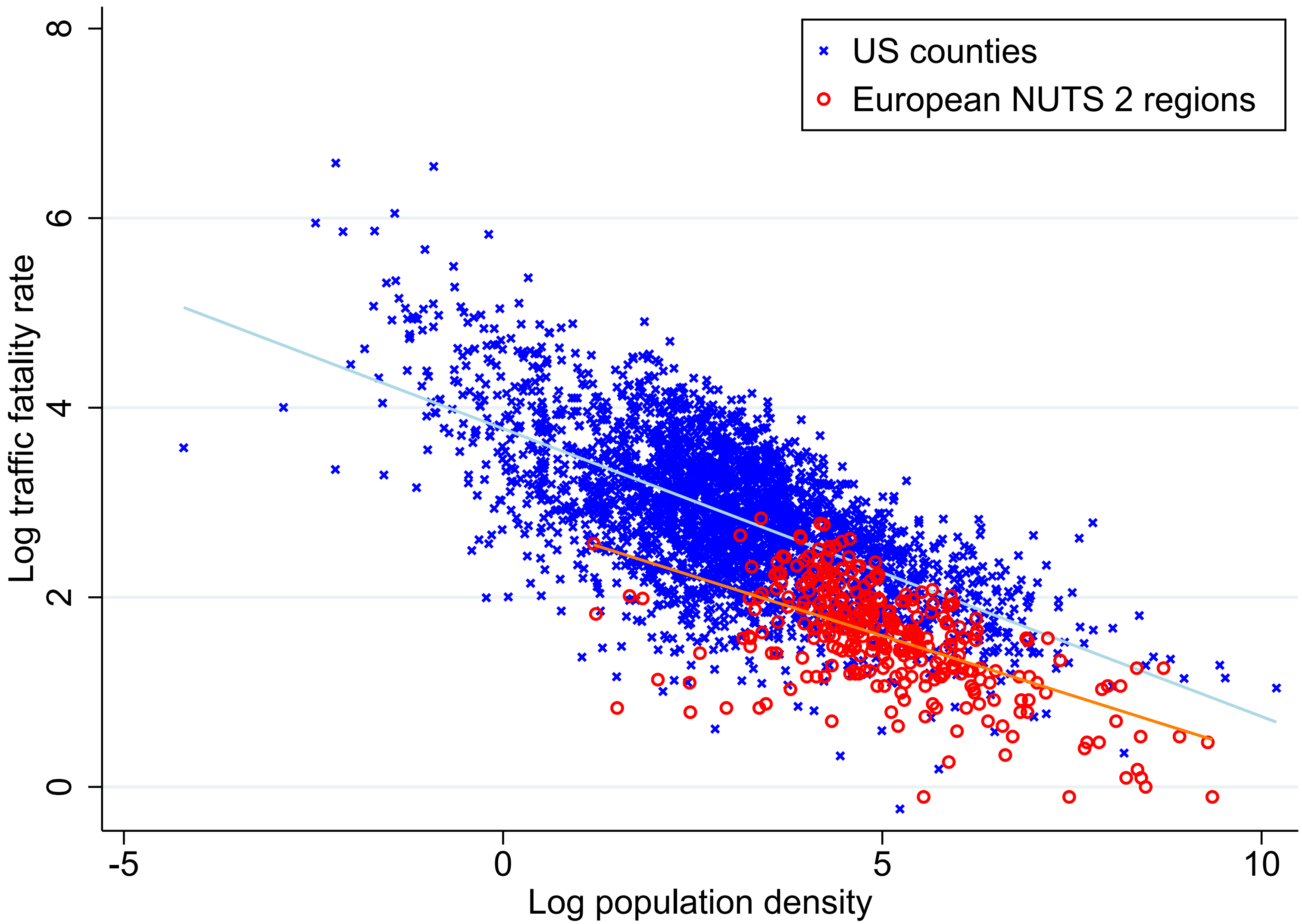

Figure 8: Comparing the U.S. and Europe.

A Online Appendix

A.1 Model Derivations

A.1.1 Finding $Z(y_i, \tau_i, L_i)$

Consider an equilibrium in which some drive and some do not drive, so that $z_i < 1$. In this equilibrium there must exist some agent who is indifferent between driving and not driving. Let that agent's driving preference, η_j, be denoted $\widetilde{\eta}$. Equalizing the expressions for V_{ij}^D and V_{ij}^{ND} in (7) and (8), with $\eta_j = \widetilde{\eta}$, gives

$$\left[\frac{(1-\tau_i)y_i - D}{(1-\tau_i)y_i}\right]^{\alpha} = \left[\frac{(\tau_i y_i)^{\lambda} L_i^{1+\lambda} z_i}{X\widetilde{\eta}}\right]^{1-\alpha}. \tag{A.1}$$

The agents who choose not to drive are those with the lowest levels of η_j. Thus, the fraction of the population with η_j below $\widetilde{\eta}$ (the level of the marginal agent) must equal $1 - z_i$, i.e., the fraction of the population who does not drive. Using the Pareto distribution in (2), it follows that

$$F(\widetilde{\eta}) = 1 - \widetilde{\eta}^{-\frac{1}{\delta}} = 1 - z_i, \tag{A.2}$$

implying

$$\widetilde{\eta} = z_i^{-\delta}. \tag{A.3}$$

Using (A.1) and (A.3) to solve for z_i gives the expression for $Z(y_i, \tau_i, L_i)$ in (18). Thus, if an equilibrium exists where some arents are not driving, then it must be such that $z_i = Z(y_i, \tau_i, L_i)$.

A.1.2 Finding τ_i^*

Next I derive τ_i^*. Recall that this is the tax rate in an equilibrium where some do not drive and drivers are in majority, $z_i \in [1/2, 1)$. Let $\widetilde{V}_i^D$ denote the utility of the agent who is indifferent between driving and not driving. Setting $\eta_j = \widetilde{\eta}$ in (7) (where, recall, $\widetilde{\eta}$ is the η_j of the marginal driver) gives

$$\widetilde{V}_i^D = \alpha \ln\left[(1-\tau_i)y_i - D\right] + (1-\alpha)\ln\left(\frac{X}{z_i L_i}\right) + (1-\alpha)\ln(\widetilde{\eta}). \tag{A.4}$$

The utility of an other agent who drives (i.e., with $\eta_j > \widetilde{\eta}$) is also given by (7). Using (A.3) and (A.4), this can be rewritten as

$$\begin{aligned} V_{ij}^D &= \alpha \ln\left[(1-\tau_i)y_i - D\right] + (1-\alpha)\ln\left(\frac{X}{z_i L_i}\right) + (1-\alpha)\ln(\eta_j) \\ &= \widetilde{V}_i^D + (1-\alpha)\ln(\eta_j) - (1-\alpha)\ln(\widetilde{\eta}) \\ &= \widetilde{V}_i^D + (1-\alpha)\ln(\eta_j) + (1-\alpha)\delta\ln(z_i) \\ &= \alpha\ln\left[(1-\tau_i)y_i\right] + (1-\alpha)\lambda\ln\left(\tau_i y_i L_i\right) + (1-\alpha)\ln(\eta_j) + (1-\alpha)\delta\ln(z_i), \end{aligned} \tag{A.5}$$

where the last equality uses the fact $\widetilde{V}_i^D = V_{ij}^{ND}$, which must hold by the definition of $\widetilde{V}_i^D$, and the expression for V_{ij}^D in (8).

The tax rate in an equilibrium where drivers are in a majority maximizes (A.5), subject to $z_i = Z(y_i, \tau_i, L_i)$. Substituting $z_i = Z(y_i, \tau_i, L_i)$ in (18) into (A.5), and ignoring terms that do not involve τ_i (containing the logs of y_i, L_i, and η_j) gives this objective function:

$$\begin{aligned} W_{ij}^D = \; & \alpha\ln(1-\tau_i) + (1-\alpha)\lambda\ln(\tau_i) \\ & +(1-\alpha)\delta\left\{\left[\tfrac{\alpha}{(1-\alpha)(1+\delta)}\right]\ln\left[\tfrac{(1-\tau_i)y_i - D}{1-\tau_i}\right] - \left(\tfrac{\lambda}{1+\delta}\right)\ln(\tau_i)\right\}. \end{aligned} \tag{A.6}$$

Rearranging, and factoring out $1/(1+\delta)$, we can write

$$W_{ij}^D = \frac{1}{1+\delta}\left\{\alpha\ln(1-\tau_i) + (1-\alpha)\lambda\ln\left(\tau_i\right) + \alpha\delta\ln\left(\left[1-\tau_i\right]y_i - D\right)\right\}. \tag{A.7}$$

Maximizing (A.7) with respect to τ_i, the first-order condition can be written

$$-\frac{\alpha}{1-\tau_i} + \frac{(1-\alpha)\lambda}{\tau_i} - \frac{\alpha\delta y_i}{(1-\tau_i)\,y_i - D} = 0. \tag{A.8}$$

The left-hand side of (A.8) is simply the function $G(\tau_i)$ in (9). Thus, the equilibrium tax rate when drivers are in a majority, τ_i^*, is defined from $G(\tau_i^*) \equiv 0$.

A.1.3 Finding τ_i^{**}

Next I derive τ_i^{**}, which is the tax rate in an equilibrium where some do not drive and non-drivers are in majority, $z_i < 1/2$. This is found by simply maximizing V_{ij}^{ND} in (8) with respect to τ_i, which gives τ_i^{**} as in (10). It can also be found by setting $\delta = 0$ in (A.8). That is, when heterogeneity is removed all agents are indifferent between driving and no driving, and thus agree on the same tax rate.

A.1.4 Finding $\widehat{L}_i$

Let $V_{ij}^{D,0}$ denote the level of V_{ij}^{D} associated with zero taxes, and thus no public transit and all agents driving. Setting $\tau_i = 0$ and $z_i = 1$ in (7) gives

$$V_{ij}^{D,0} = \alpha \ln (y_i - D) + (1-\alpha) \ln \left(\frac{\eta_j X}{L_i} \right). \tag{A.9}$$

Similarly, let V_{ij}^{D,τ^*} denote the level of V_{ij}^{D} associated with the optimal tax rate when drivers are in majority and some do not drive, i.e., $\tau_i = \tau_i^*$ and $z_i < 1$. From (7), this can be written

$$V_{ij}^{D,\tau^*} = \alpha \ln \left[(1-\tau_i^*) y_i - D \right] + (1-\alpha) \ln \left(\frac{\eta_j X}{z_i L_i} \right). \tag{A.10}$$

Setting $V_{ij}^{D,\tau^*} = V_{ij}^{D,0}$ in (A.9) and (A.10) we can solve for what z_i must equal for agents to be indifferent between the two tax rates ($\tau_i = \tau_i^*$ and $\tau_i = 0$, respectively). Denote that level of z_i by z_i^*. Equalizing (A.9) and (A.10), setting $z_i = z_i^*$, gives the expression in (13).

The next task is to find what level of population implements $z_i = z_i^*$. Using (18), with $\tau_i = \tau_i^*$, gives

$$z_i^* = Z(y_i, \tau_i^*, L_i) = \left[\frac{(1-\tau_i^*) y_i - D}{(1-\tau_i^*) y_i} \right]^{\frac{1}{(1-\alpha)(1+\delta)}} \left[\frac{X}{(\tau_i^* y_i)^{\lambda}} \right]^{\frac{1}{1+\delta}} L_i^{-\frac{1+\delta}{1+\lambda}} = B_i^* L_i^{-\frac{1+\delta}{1+\lambda}}, \tag{A.11}$$

where $B_i^* = Z(y_i, \tau_i^*, 1)$ is defined in (12). Solving (A.11) for L_i gives $\widehat{L}_i$ in (11).

A.1.5 Finding $\widehat{\widehat{L}}_i$

Recall that $\widehat{\widehat{L}}_i$ is the population threshold above which non-drivers become a majority, implementing their preferred tax rate, τ_i^{**}. Thus, $\widehat{\widehat{L}}_i$ is defined from $Z(y_i, \tau_i^{**}, \widehat{\widehat{L}}_i) = 1/2$. Using (18), with $\tau_i = \tau_i^{**}$, gives

$$\widehat{\widehat{L}}_i = \left\{ 2 \left[\frac{(1-\tau_i^{**}) y_i - D}{(1-\tau_i^{**}) y_i} \right]^{\frac{1}{(1-\alpha)(1+\delta)}} \left[\frac{X}{(\tau_i^{**} y_i)^{\lambda}} \right]^{\frac{1}{1+\delta}} \right\}^{\frac{1+\delta}{1+\lambda}}, \tag{A.12}$$

which can be rewritten as in (14) and (15).

A.2 Allowing for Migration

In the main text I let L_i vary exogenously. This section explores migration by letting L_i adjust endogenously to equalize utilities across locations.

To that end, I first assume away heterogeneity between agents, setting $\delta = 0$ and thus $\eta_j = 1$ for all agents j. That way I do not need to keep track of an endogenous type distribution across locations, as different types may migrate to different locations. The main implication of this assumption is that $\delta > \lambda$ cannot hold, meaning the model can no longer produce a positive relationship between $z_i L_i$ and L_i. Thus, I here also set $\lambda = 1$ to simplify the notation. This means that public transit provision is here linear in tax revenues; cf. (5).

With $\delta = 0$ and $\lambda = 1$ in (9) and (10), it is straightforward to see that the two tax rates are identical:

$$\tau_i^* = \tau_i^{**} = 1 - \alpha. \tag{A.13}$$

Similarly, we see from (12) and (15) that

$$B_i^* = B_i^{**} = \left[\frac{\alpha y_i - D}{\alpha y_i}\right]^{\frac{\alpha}{1-\alpha}} \frac{X}{(1-\alpha)y_i}. \tag{A.14}$$

and

$$z_i^* = \left[\frac{\alpha y_i - D}{y_i - D}\right]^{\frac{\alpha}{1-\alpha}}. \tag{A.15}$$

With $(1+\delta)/(1+\lambda) = 1/2$, it then follows that $\widehat{L}_i$ in (11) becomes

$$\widehat{L}_i = \left(\frac{B_i^*}{z_i^*}\right)^{\frac{1}{2}} = \left[\frac{(y_i - D)^\alpha}{\alpha^\alpha (1-\alpha)^{1-\alpha} y_i}\right]^{\frac{1}{2(1-\alpha)}} X^{\frac{1}{2}}. \tag{A.16}$$

We can now write the tax rate as

$$\tau_i = \begin{cases} 0 & \text{if } L_i \leq \widehat{L}_i, \\ \\ 1-\alpha & \text{if } L_i > \widehat{L}_i, \end{cases} \tag{A.17}$$

and the equilibrium fraction driving as

$$z_i = \begin{cases} 1 & \text{if } L_i \leq \widehat{L}_i, \\ \\ \left(\frac{\alpha y_i - D}{\alpha y_i}\right)^{\frac{\alpha}{1-\alpha}} \frac{X}{(1-\alpha) y_i L_i^2} & \text{if } L_i > \widehat{L}_i, \end{cases} \tag{A.18}$$

where we use (A.17) and (18) with $\delta = 0$ and $\lambda = 1$.

As in the model without migration, when L_i exceeds $\widehat{L}_i$ it becomes optimal to switch from a regime with no taxes and only driving to one with taxation and some public transit provision. Note, however, that in this setting we do not need to know whether drivers or non-drivers are in a majority, as long as some agents use public transit. The reason is that agents are ex-ante identical and thus indifferent between the two modes of transportation in equilibrium, and therefore prefer the same tax policy.

To explore migration decisions, I derive payoffs to living in location i. Substituting the expressions for τ_i and z_i in (A.17) and (A.18) into the utility function defined by (6) to (8), with $\eta_j = 1$, gives us this indirect utility function:

$$U_i = \max\{U_i^D, U_i^{ND}\} = \begin{cases} U_i^D & \text{if } L_i \leq \widehat{L}_i, \\ \\ U_i^{ND} & \text{if } L_i > \widehat{L}_i, \end{cases} \tag{A.19}$$

where U_i^D and U_i^{ND} are given by

$$U_i^D = \alpha \ln(y_i - D) + (1-\alpha)\ln\left(\frac{X}{L_i}\right), \tag{A.20}$$

and

$$U_i^{ND} = \ln\left[\alpha^\alpha (1-\alpha)^{1-\alpha}\right] + \ln(y_i) + (1-\alpha)\ln(L_i). \tag{A.21}$$

We can interpret U_i^D and U_i^{ND} as the utilities associated with the two tax regimes ($\tau_i = 0$ and $\tau_i = 1-\alpha$, respectively) being imposed exogenously.

Figure A.2 plots various variables against L_i in three stacked panels: U_i^D and U_i^{ND} in the top panel; $U_i = \max\{U_i^D, U_i^{ND}\}$ in the middle panel; and z_i in the bottom panel. The relationship between U_i and L_i is U-shaped, with U_i minimized at $L_i = \widehat{L}_i$, where the shift to public transit occurs. For $L_i \leq \widehat{L}_i$, everyone is driving ($z_i = 1$), and $U_i = U_i^D$ is decreasing in L_i. This is due to the increased congestion caused by adding more drivers as population expands, reducing transportation consumption.

For $L_i > \widehat{L}_i$, the fraction driving falls below one and is decreasing in L_i, while $U_i = U_i^{ND}$ is increasing in L_i. Here larger population brings more tax revenue with which to fund public transit provision, which raises utility for those that use public transit, and thus also drivers, since agents are indifferent between driving and using public transit.

A.2.1 A Two-Location Setting: The Symmetric Case

I next consider a case with two locations, indexed $i \in \{A, B\}$, with populations L_A and L_B, such that total population across both locations sums up to unity, $L_A + L_B = 1$.

Agents earn income y_A (y_B) if they choose to live in location A (B). For now, I consider the fully symmetric case, letting incomes be the same in two locations, here denoted just y.

With free (or frictionless) migration, the associated indirect utilities, U_A and U_B, must equalize in equilibrium. This is illustrated in Figure A.3, which has two panels, both of which show the population in location A on the horizontal axis. We can thus read the population in location B as $L_B = 1 - L_A$. The top panel plots the indirect utilities (U_A and U_B), corresponding to that in the middle panel in Figure A.2, while the bottom panel of Figure A.3 shows car dependency in the two locations (z_A and z_B).

Since the two locations are symmetric with respect to their exogenous variables, they face the same population thresholds for using public transit. That threshold is here assumed to fall below 1/2, i.e.,

$$\widehat{L}_A = \widehat{L}_B = \left[\frac{(y-D)^{\alpha}}{\alpha^{\alpha}(1-\alpha)^{1-\alpha} y} \right]^{\frac{1}{2(1-\alpha)}} X^{\frac{1}{2}} < 1/2, \tag{A.22}$$

which holds for small enough X.

Under this assumption, three equilibria exist:

1. A locally unstable equilibrium with population of 1/2 in each location.

2. A locally stable equilibrium with populations $\overline{L}_A \in (1/2, 1)$ in location A, and $\underline{L}_B \in (0, 1/2)$ in location B, such that $\underline{L}_B + \overline{L}_A = 1$.

3. A locally stable equilibrium with populations $\overline{L}_B \in (1/2, 1)$ in location B, and $\underline{L}_A \in (0, 1/2)$ in location A, such that $\underline{L}_A + \overline{L}_B = 1$.

Proofs and derivations and deferred to Section A.2.3 below.

The equilibria of type 2 and 3 are simply each other's mirror images, i.e., $\underline{L}_A = \underline{L}_B$ and $\overline{L}_A = \overline{L}_B$. The equilibrium of type 1 is not very interesting, since it is unstable.

A.2.2 A Two-Location Setting: The Asymmetric Case

The case where the two locations differ in income levels can be understood from Figure A.4, which is otherwise numerically identical to that in Figure A.3, but such that $y_B > y_A$. This means that the curve for indirect utility in location B shifts up relative to that for location A. As long as the shift is small enough the equilibrium where location A has larger population than location B, and lower car dependency, still exists. That is, when comparing the two locations the poorer one can have larger population but less car dependency.

This mirrors the result in Section 4.3.4. That is, the relationship between population density and car dependency (L_i and z_i) is here negative, while that between per-capita incomes and car dependency (y_i and z_i) can be positive or negative.

A.2.3 Finding and Characterizing the Equilibria

First I show that the **equilibrium of type 1** exists, which amounts to showing that $U_A = U_B$ when the population equals 1/2 in each location. We have assumed that $\widehat{L}_A = \widehat{L}_B < 1/2$, so we know that $z_i < 1$ and $U_i = U_i^{ND}$ for $i \in \{A, B\}$. Setting $L_A = L_B = 1/2$ in (A.21), with $y_i = y$, we get $U_A = U_B$.

To show that this equilibrium is locally unstable, set $L_A = 1/2 + \varepsilon$, and $L_B = 1/2 - \varepsilon$, for some small $\varepsilon > 0$. Since (A.21) is increasing in L_i, this implies $U_A > U_B$. Vice versa, setting $L_A = 1/2 - \varepsilon$, and $L_B = 1/2 + \varepsilon$, for some small $\varepsilon > 0$, gives $U_B > U_A$.

Next I show that an **equilibrium of type 2** exists. We need to show that there exist $\overline{L}_A > 1/2$ and $\underline{L}_B < 1/2$, such that $\underline{L}_B + \overline{L}_A = 1$ and $U_A = U_B$ when $L_A = \overline{L}_A$ and $L_B = \underline{L}_B$. Since $\overline{L}_A > 1/2$, it must hold in equilibrium that

$$U_A = U_A^{ND} = \ln\left[\alpha^\alpha (1-\alpha)^{1-\alpha}\right] + \ln(y) + (1-\alpha)\ln(\overline{L}_A),$$

which implies $U_B = U_B^D$, since otherwise (i.e., if $U_B = U_B^{ND}$) we must have $\underline{L}_B = \overline{L}_A$, which cannot hold, since we postulated that $\underline{L}_B < 1/2 < \overline{L}_A$; thus, this equilibrium (if it exists) must be such that

$$U_B = U_B^D = \alpha \ln(y - D) + (1-\alpha)\ln\left(\frac{X}{1-\overline{L}_A}\right).$$

Write the difference between $U_A = U_A^{ND}$ and $U_B = U_B^{D}$ for any level of L_A as

$$\begin{aligned}\Delta = \;& \ln\left[\alpha^{\alpha}(1-\alpha)^{1-\alpha}\right] + \ln(y) + (1-\alpha)\ln(L_A) \\ & -\alpha \ln(y-D) + (1-\alpha)\ln\left(\frac{X}{1-L_A}\right).\end{aligned}$$

We can now search for a level of L_A on the interval $(1-\widehat{L}_A, 1) = (1-\widehat{L}_B, 1)$, such that $\Delta = 0$; cf. Figure A.3.

Consider first the point $L_A = 1-\widehat{L}_A = 1-\widehat{L}_B$, where $L_B = 1 - L_A = \widehat{L}_B$. Here U_B reaches its minimum point, but U_A does not (due to symmetry and $L_A \neq \widehat{L}_A = \widehat{L}_B$). Thus $\Delta > 0$.

Next we note that $\Delta \to -\infty$ as $L_A \to 1$ (and $L_B \to 0$).

Thus, there must exist some level of L_A on the interval $(1-\widehat{L}_A, 1)$, such that $\Delta = U_A - U_B = 0$. That level is $\overline{L}_A$. Since $1-\widehat{L}_A > 1/2$, it follows that $\overline{L}_A \in (1/2, 1)$.

To show that this equilibrium is locally stable, note that

$$\frac{\partial \Delta}{\partial L_A} = (1-\alpha)\left(\frac{1}{L_A} - \frac{1}{1-L_A}\right) = \frac{(1-\alpha)(1-2L_A)}{L_A(1-L_A)},$$

which implies $\partial\Delta/\partial L_A < 0$ at $L_A = \overline{L}_A > 1/2$.

To show that the **equilibrium of type 3** exists and is stable follows the same steps as those for the equilibrium of type 2, except that A and B switch places.

A.3 Data

A.3.1 United States Aggregate Data

The analysis in Section 3 relies on two sources. First I use data from the National Safety Council, a self-described U.S. non-profit group advocating for traffic safety. From that website I followed the links "Data Table" and "Download Excel Sheet" to download a file called "Motor-Vehicle Deaths and Rates.xlsx." This file contains time-series data on traffic fatality rates per 100,000 population (in column H of the .xlsx file), and on the aggregate number of vehicles (column C).

Then I downloaded population data from USA Facts, another non-profit. On that website I clicked on "Download Data" to access a file called population_usafacts.csv, which contains data in total U.S. population from 1900. Merging this .csv file with the .xlsx file described above allows me to calculate the vehicle ownership rate per capita as the aggregate number of vehicles over total U.S. population.

A.3.2 United States County Data

Traffic fatalities: Data on traffic fatalities are from the Fatality and Injury Reporting System Tool (FIRST), an online data portal set up by the National Highway Traffic Safety Administration (NHTSA). To retrieve the data, I first select People from the options Crashes, Vehicles, People, Drivers, Occupants, Pedestrians, and Pedalcyclists. Under Select Fatality and/or Injury, I choose Persons Killed in Fatal Crashes. I then set Select Time Frame to 2010, which is the year for which I have (good) data on population density; see below. I leave blank under Select State or Region and Filter Your Selection. Finally, under Build Your Report, I enter the data elements State and County for Rows, and Crash Date (year) for Columns. When clicking Submit a new windows opens. Scrolling down I can download an Excel file (called CrashReport). This contains total traffic deaths by county. Using data on total population by county, as described below, I then calculate the traffic fatality rate per $100,000$ population.

Population density and urbanization. Data on population density and urbanization come from the website 2010 Census Urban and Rural Classification and Urban Area Criteria, hosted by the U.S. Census Bureau. The data I use is in the Excel file "Percent Urban and Rural In 2010 by State and County [$<$ 1.0 MB]," downloadable from the website. To calculate population density I divide total population (POP_COU) by area (AREA_COU), and then multiply by $1,000,000$ to convert from square meters to square kilometers. The urbanization rate is computed as urban population (POP_URBAN) divided by total population (POP_COU).

Per-capita income: Per-capita income data are from the Bureau of Economic Analysis and refer to per capita personal income (in dollars) for the year 2010. The data were retrieved by first selecting "County and MSA personal income summary: personal income, population, per capita personal income" under the tab Table. I then make the following selections: County under Major Area; "All counties is the US" under State; "Per capita personal income" under Statistic; Levels under Unit of Measure; and 2010 under Time Period.

Life expectancy: Data on life expectancies at birth were downloaded from the Data Visualization Gallery hosted by the National Center for Health Statistics at the Centers for Disease Control and Prevention. The download was made through a link labelled "Download Datasets: CSV Format," and the variable used is in a column labelled just "Life Expectancy" (with other columns reporting range and standard errors). The data refer to life expectancies at birth 2010-2015 and are provided as averages at the census tract level, finer than counties. They were aggregated to the county level by taking the (unweighted) average across census tracts in each county. This dataset does not provide numerical identifiers (County FIPS Codes) so the merging with other data was done by county and state name in the format "Autauga, AL."

Homicide rates: Data on homicide (murder) rates are from the Compressed Mortality database, hosted by the Centers for Disease Control and Prevention. To retrieve the full dataset, I follow the web link and click Agree to the conditions, after which a new webpage opens providing a data portal. There I make the following selections:

- Under *1. Organize table layout*, I select "Group Results By County And By Year."
- Under *2. Select location, State/County or CBSA*, I indicate "States" under "Click a button to choose locations by State, Census Region, or HHS Region." Then I indicate "*All* (The United States)" under Browse and States.
- Under *2.a. Select urbanization classifications*, I select 2013 and All Categories.
- Under *3. Select years and demographics*, I select "All Ages," "All Genders," "All Races," and "All Origins." Under Year I highlight all years from 1999 to 2016; I do not highlight "All Years."
- Under *4. Select cause of death*, I select the option "Injury Intent and Mechanism" and the indicate "Homicide" under "Injury Intent" and "All Causes of Death" under "Injury Mechanism & All Other Leading Causes."

- Under *5. Other Options*, I indicate "Show Totals," "Show Zero Values," and "Show Suppressed Values."

After clicking Send a new webpage opens, from where the data was downloaded in .txt format by clicking Export. The downloaded file was then converted into .xlsx format to be read in Stata.

The homicide rate per 100,000 population is calculated as the number of deaths over reported population (divided by 100,000) for each county and year. In the downloaded data these variables are labelled Deaths and Population. I then calculate the mean homicide rate for each county across all years from 2006 to 2016 (2016 being the last year with data reported). These means may be based on different sets of years for different counties, since many county-years have missing data. Choosing the same year for all counties (e.g., 2009 or 2010) shrinks the sample further.

Historical population density: Data on historical population levels come from a website maintained by Andrew J. Van Leuven (Van Leuven, 2020), also posted on GitHub. The data can be downloaded in .csv format through the link "Here is a direct download of the data." For county area I use the variable (AREA_COU), retrieved from the U.S. Census Bureau to calculate modern population density (see above). To compute population density I divide Van Leuven's population measure for 1900 (pop_1900) by the county's land area, and then multiply by $1,000,000$ to convert from square meters to square kilometers.

Republican vote share and property taxes: The political variables are from Bazzi et al. (2020) and was downloaded in Stata (.dta) format from Wiley's Online Library. The data come in a zip folder, found under "Supporting Information" and "Data and Programs." The filename is ecta200214-sup-0002-dataandprograms.zip and the dataset is called proptaxvote.dta, found in the unzipped folder "data." The variables used here are called avgrep2000to2016 (the Republican Presidential vote share 2000-2016) and propertytaxrate2010 (the property tax rate in 2010).

The fraction commuters using public transit: This variable is based on numbers from the American Community Survey, conducted by the U.S. Census Bureau and referring to the year 2010. The table is called B08130–MEANS OF TRANSPORTATION TO WORK BY PLACE OF WORK–STATE AND COUNTY LEVEL, which can be downloaded as a zipped file, and unzipped as a .csv file. I first calculated the total number of commuters as the sum of all people using public transit, those driving alone, and those car pooling. The fraction using public transit is then calculated as the total number of people using public transit over the total number of commuters.

A.3.3 European Regional Data

All data for Europe (and some neighboring countries) come from the Eurostat database linked to here. This is the body of the European Union that collects and maintains data from the nationals statistical agencies of both EU member countries, and a few other countries, such as Turkey and some European Free Trade Association (EFTA) countries.

The statistical classification system known as NUTS (Nomenclature of Territorial Units for Statistics) has three levels of subnational divisions. I here use the intermediate level (NUTS 2), which permits access to data to all of the variables I am interested in, in particular traffic deaths, which I have not been able to access at the more disaggregate NUTS 3 level.

Since NUTS 2 regions are larger than U.S. counties, I use GDP per capita for NUTS 2 regions and per-capita incomes for U.S. counties.

General first steps: For all variables below, the process involves first searching under a specific code in the Eurostat database linked to above, which gives access to various datasets containing different variables. After clicking "Access dataset," a portal opens where I can make selections. In all cases, under the tab Selection I set the available geopolitical units for rows, and years (or Time) for columns. I also click the tab Format to indicate "codes and labels," which ensures that the downloaded data contains both region names and the relevant code (for later merging). Under the tab Download I choose "Full dataset [with the code for dataset indicated]," and then "Spreadsheet (.xlsx)." This downloads a file in .xlsx format, which can cleaned in Excel and/or be read directly in Stata.

From that file, I choose the year 2018 (the latest with UK data) for all variables below, unless otherwise indicated. The main exception is homicide rates, where I use the latest year available, 2010.

Population density: Data on population density is accessed by searching under the code tgs00024 in the Eurostat database. This allows access to a dataset called "Population density by NUTS 2 region." The download produces an .xlsx file with a name containing tgs00024. The numbers are report as persons per square kilometre, so no further computations are needed after reading the data in Stata.

Urbanization: The urbanization data can be retrieved by searching under the code lfst_r_lfsd2hh in the Eurostat database. Under the tab "Degree of urbanisation [deg_urb]," I can click "Check All" to highlight "Cities," "Towns and suburbs," "Rural areas," and "No response," The download produces an .xlsx file containing fst_r_lfsd2hh, and a few sheets. The number of city households (coded as DEG1 in the spreadsheet) is reported in Sheet 2, measured in thousands. The total number of households (coded TOTAL) is reported in Sheet 1, also in thousands. The urbanization rate is calculated as the number of city households over the total number of households.

GDP per capita: Data on Gross Domestic Product per capita (or per inhabitant) can be found by searching under the code tgs00005 in the Eurostat database. The downloaded .xlsx file containing the code tgs00005. The GDP per capita measurements are corrected for purchasing power differences, or what in the downloaded file is called "Purchasing power standard (PPS, EU27 from 2020), per inhabitant [PPS_EU27_2020_HAB]."

Traffic fatalities: Data on the total number of fatalities in traffic can be found be searching under tran_r_acci in the Eurostat database. I adjust the selections to "Check All" under the tabs "Type of victim" and "Unit of measure [unit]." This ensures that the download includes data on the total number killed per million inhabitants (as well as totals). The downloaded .xlsx file contains tran_r_acci, and the variable used in the analysis can be found in Sheet 2 of the downloaded file. To get the rate in units per 100,000 people, rather than per million (the reported format), I multiply the reported values by 10.

Life expectancy: To download data on life expectancy at birth by (sex and) NUTS 2 region, I searched under the code tgs00101 in the Eurostat database. I then selected "Check All" under the tab Sex, meaning data are reported separately for male, female, and total population. The downloaded file contains the code tgs00101. The variable used here are for total population, reported in Sheet 1 of the file, with these settings.

Total population: Data on total population can be found be searching under the code tgs00096 in the Eurostat database linked to above. This allows download of an .xlsx file with the code tgs00096, which reports population on January 1, for all years selected, by NUTS 2 region. From the downloaded file, I select the years 2010 and 2018 (both referring to January 1) because these population totals are used to compute homicide rates and the number vehicles per capita below.

Homicides: To compute homicide rates I first retrieve data on homicide totals by searching under the code crim_gen_reg in the Eurostat database. This gives me access to a dataset called "Crimes recorded by the police by NUTS 3 regions," which also includes NUTS 2 level data. I select "Check All" under the tab "International classification of crime for statistical purposes (ICCS)," which makes the download include a couple of different types of crime, of which the focus here is on "Intentional homicide." With these settings, the downloaded .xlsx file contains the code crim_gen_reg, with homicide totals reported in Sheet 1 of the downloaded file.

At the time I made the download, the homicide (and other crime) totals are available only for the years 2008-2010, from which I selected 2010. To get the homicide rate I divide total homicides by total population in 2010 (see above) after matching by NUTS 2 codes in Stata. I then multiply by 100,000 to get the rate as homicides per 100,000 population.

Note also, as already mentioned, that these data are provided at both the NUTS 2 and NUTS 3 levels, the latter being more disaggregated, but here I use only observations at the NUTS 2 level.

Vehicles per capita: To retrieve data on the number of registered vehicles I search under the code tran_r_vehst in the Eurostat database (link above). This produces a link to access a dataset called "Stock of vehicles by category and NUTS 2 regions." Under the tab Vehicles I indicate "Check All," so that the download includes data on various types of vehicles. Under the tab "Unit of measure" I indicate only "Number," and not units per thousand inhabitants, because the latter has many missing observations. With these settings, the downloaded .xlsx file contains the code tran_r_vehst, with total number of registered vehicles, trailers and motorcycles, reported in Sheet 1 of the downloaded file; the variable is called "All vehicles (except trailers and motorcycles) [TOT_X_TM]." From this sheet I select the column with year 2018, and to get vehicles per capita I divide the total number vehicles by total population in 2018 (see above) after matching by NUTS 2 codes in Stata.

Motorways per capita: To retrieve data on the length of motorways I search under the code tran_r_net in the Eurostat database. Under the tab "Transport infrastructure," I select "Check All" to make sure that the variable "Motorways [MWAY]" is included in the download. Under the tab "Unit of measure [unit]" I select "Kilometre [KM]." The downloaded file has the code tran_r_net in the name. The year 2018 is selected. I then divide by 2018 population to get a variable measured in kilometers per capita.

Online Appendix Tables and Figures

	Dependent variable is the log traffic fatality rate				
	(1)	(2)	(3)	(4)	(5)
Log urbanization rate	-0.335^{***}	-0.213^{***}	-0.219^{***}	-0.247^{***}	-0.247^{***}
	(0.031)	(0.028)	(0.028)	(0.030)	(0.045)
Log income per capita		-1.043^{***}	-0.801^{***}	-0.786^{***}	-0.786^{***}
		(0.084)	(0.098)	(0.099)	(0.171)
Log life expectancy			-2.792^{***}	-2.115^{***}	-2.115^{**}
			(0.550)	(0.701)	(0.791)
R^2	0.12	0.20	0.21	0.31	0.31
Number of obs.	2349	2297	2228	2228	2228
Fixed effects	None	None	None	State	State
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. Robust standard errors are indicated in parentheses, except for column (5), which clusters on state. The unit of observation is a county. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A.1: Traffic Deaths and Urbanization: United States.

	Dependent variable is the log traffic fatality rate				
	(1)	(2)	(3)	(4)	(5)
Log urbanization rate	-0.347^{***}	-0.221^{***}	-0.230^{***}	-0.331^{***}	-0.331^{***}
	(0.084)	(0.068)	(0.070)	(0.068)	(0.072)
Log GDP per capita		-0.689^{***}	-0.616^{***}	-0.284^{**}	-0.284
		(0.070)	(0.086)	(0.111)	(0.178)
Log life expectancy			-1.569^{*}	0.984	0.984
			(0.839)	(2.660)	(3.513)
R^2	0.19	0.48	0.48	0.66	0.66
Number of obs.	221	221	221	221	221
Fixed effects	None	None	None	Country	Country
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. Robust standard errors are indicated in parentheses, except for column (5), which clusters on country. The unit of observation is a NUTS 2 region. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A.2: Traffic Deaths and Urbanization: Europe.

	Dependent variable is the log homicide rate				
	(1)	(2)	(3)	(4)	(5)
Log population density	−0.180***	−0.062**	−0.066***	−0.044	−0.044
	(0.026)	(0.027)	(0.023)	(0.029)	(0.044)
Log income per capita		−1.255***	−0.078	−0.055	−0.055
		(0.163)	(0.140)	(0.140)	(0.123)
Log life expectancy			−16.034***	−15.570***	−15.570***
			(1.024)	(1.415)	(1.876)
R^2	0.13	0.26	0.57	0.65	0.65
Number of obs.	477	467	436	436	436
Fixed effects	None	None	None	State	State
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. Robust standard errors are indicated in parentheses, except for column (5), which clusters on state. The unit of observation is a county. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A.3: Homicides and Population Density: United States.

	Dependent variable is the log homicide rate				
	(1)	(2)	(3)	(4)	(5)
Log population density	0.043	0.139***	0.125***	0.050	0.050
	(0.043)	(0.049)	(0.046)	(0.034)	(0.039)
Log GDP per capita		−0.912***	−0.644***	−0.063	−0.063
		(0.110)	(0.144)	(0.141)	(0.204)
Log life expectancy			−4.828***	−12.564***	−12.564***
			(1.477)	(2.738)	(3.063)
R^2	0.01	0.32	0.35	0.74	0.74
Number of obs.	160	154	153	153	153
Fixed effects	None	None	None	Country	Country
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: Ordinary least squares regressions. Robust standard errors are indicated in parentheses, except for column (5), which clusters on country. The unit of observation is a NUTS 2 region. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A.4: Homicides and Population Density: Europe.

	(1)	(2)	(3)	(4)	(5)
	2nd stage: dependent variable is the log traffic fatality rate				
Log population density (instrumented)	-0.267^{***}	-0.259^{***}	-0.284^{***}	-0.307^{***}	-0.307^{***}
	(0.011)	(0.012)	(0.013)	(0.017)	(0.019)
Log income per capita		-0.537^{***}	-0.143^{*}	-0.043	-0.043
		(0.068)	(0.079)	(0.095)	(0.096)
Log life expectancy			-4.650^{***}	-1.973^{***}	-1.973^{***}
			(0.482)	(0.584)	(0.656)
R^2	0.39	0.40	0.42	0.47	0.47
Number of obs.	2602	2566	2493	2493	2493
	1st stage: dependent variable is log population density in 2010				
Log population density 1900	0.778^{***}	0.755^{***}	0.711^{***}	0.745^{***}	0.745^{***}
	(0.022)	(0.022)	(0.022)	(0.033)	(0.105)
Log income per capita		1.649^{***}	2.094^{***}	2.459^{***}	2.459^{***}
		(0.107)	(0.138)	(0.146)	(0.187)
Log life expectancy			-4.795^{***}	-3.345^{***}	-3.345^{**}
			(0.769)	(0.871)	(1.701)
R^2	0.52	0.57	0.57	0.70	0.70
Number of obs.	2602	2566	2493	2493	2493
1st-stage *F*-statitic	319.24	261.78	285.33	160.04	29.07
p-value	0.0000	0.0000	0.0000	0.0000	0.0000
Fixed effects	None	None	None	State	State
Standard errors	Robust	Robust	Robust	Robust	Clustered

Notes: First- and second-stage IV regressions, where modern population density (in 2010) is instrumented with population density in 1900. Robust standard errors are indicated in parentheses, except for column (5), which clusters on state. The 1st-stage *F*-statistic refers to the Anderson-Rubin *F* test of significance of endogenous regressors, retrieved as e(arf) after using the command ivreg2 in Stata. The unit of observation is a county. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A.5: Traffic Deaths and Modern Population Density Instrumented With Population Density in 1900: United States.

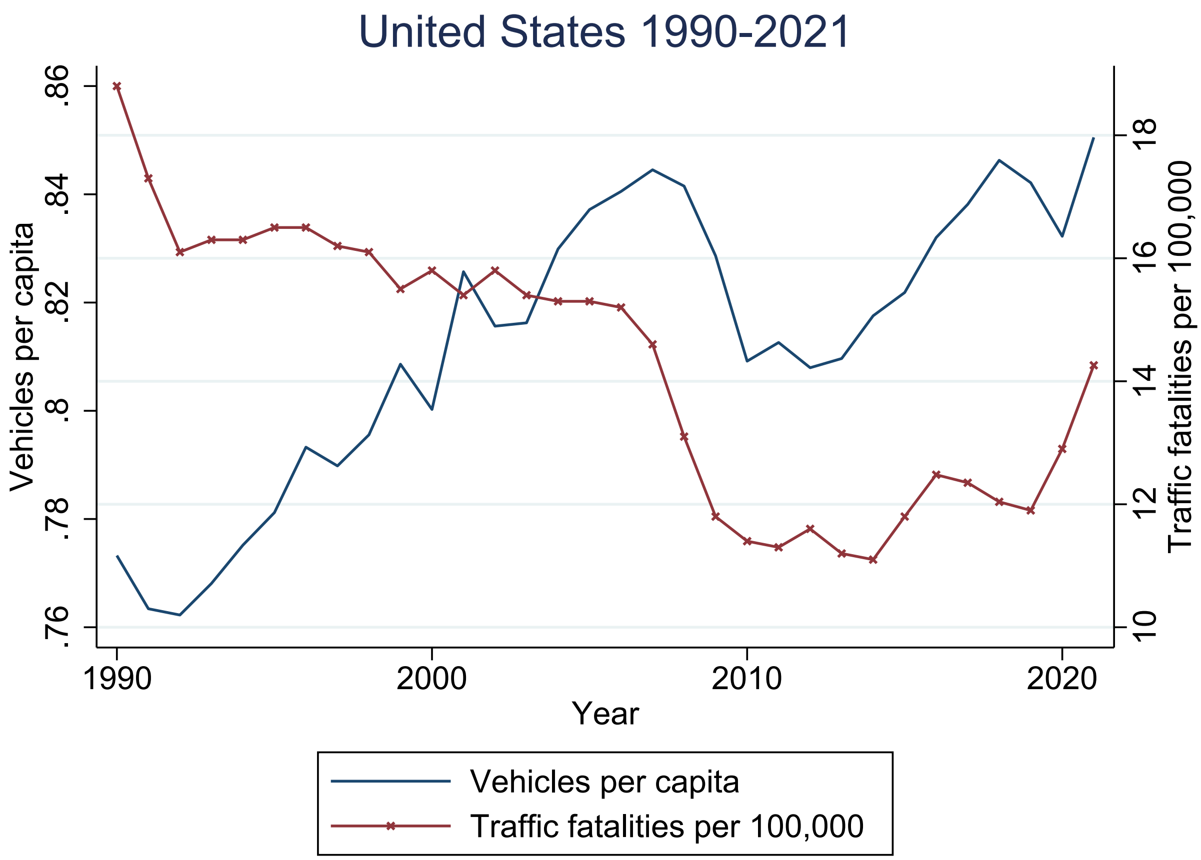

Figure A.1: Aggregate time trends in the United States from 1990.

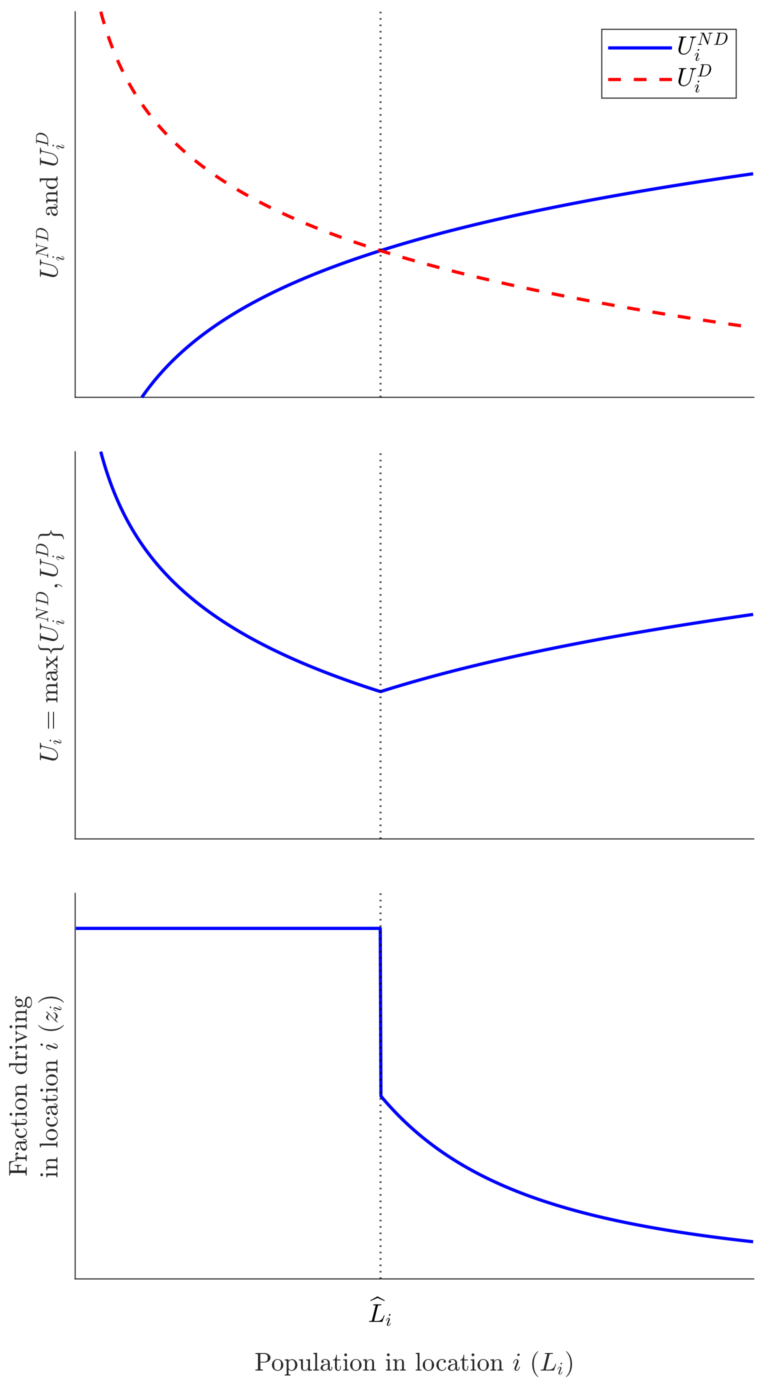

Figure A.2: Equilibrium in one location.

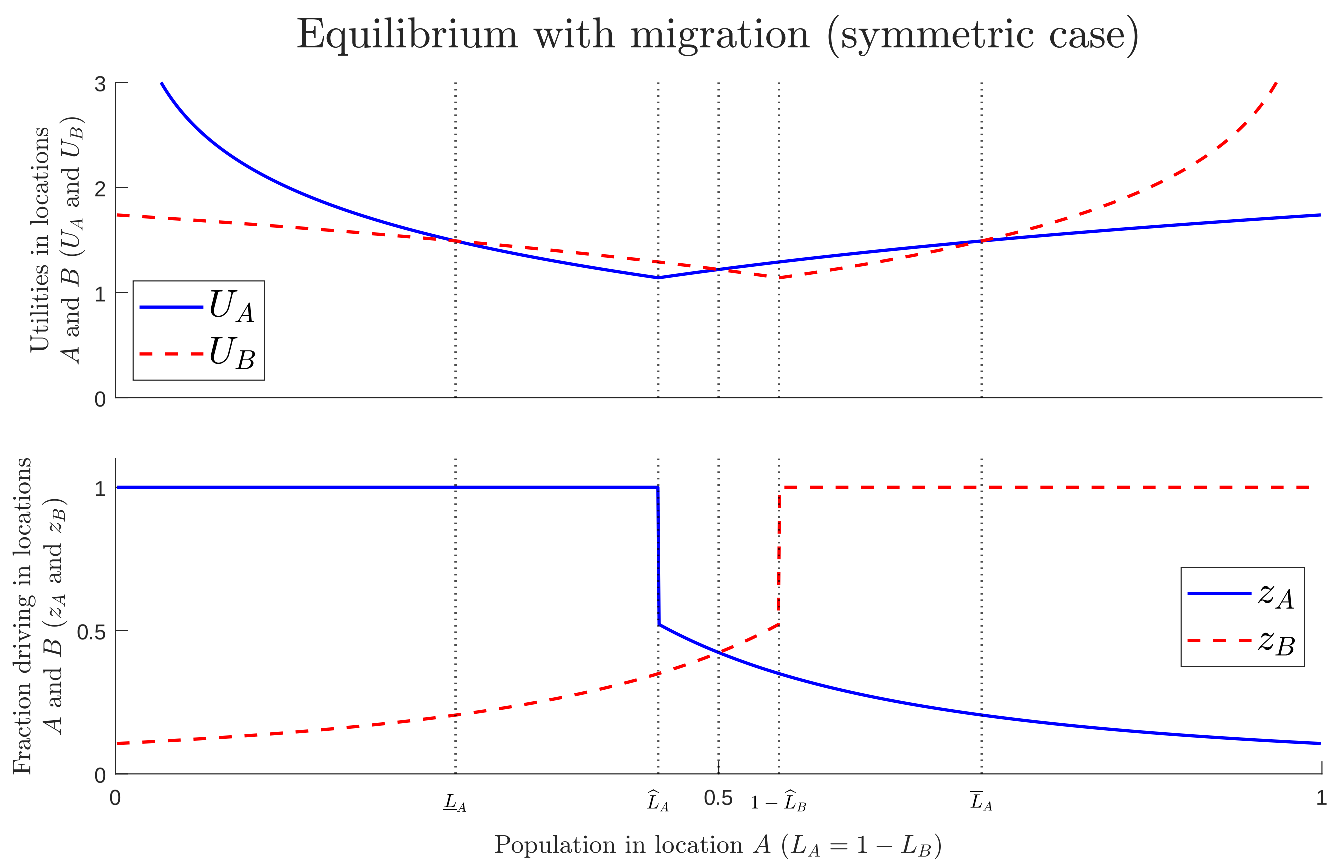

Figure A.3: Equilibrium with migration between two locations, A and B, when $y_B = y_A$.

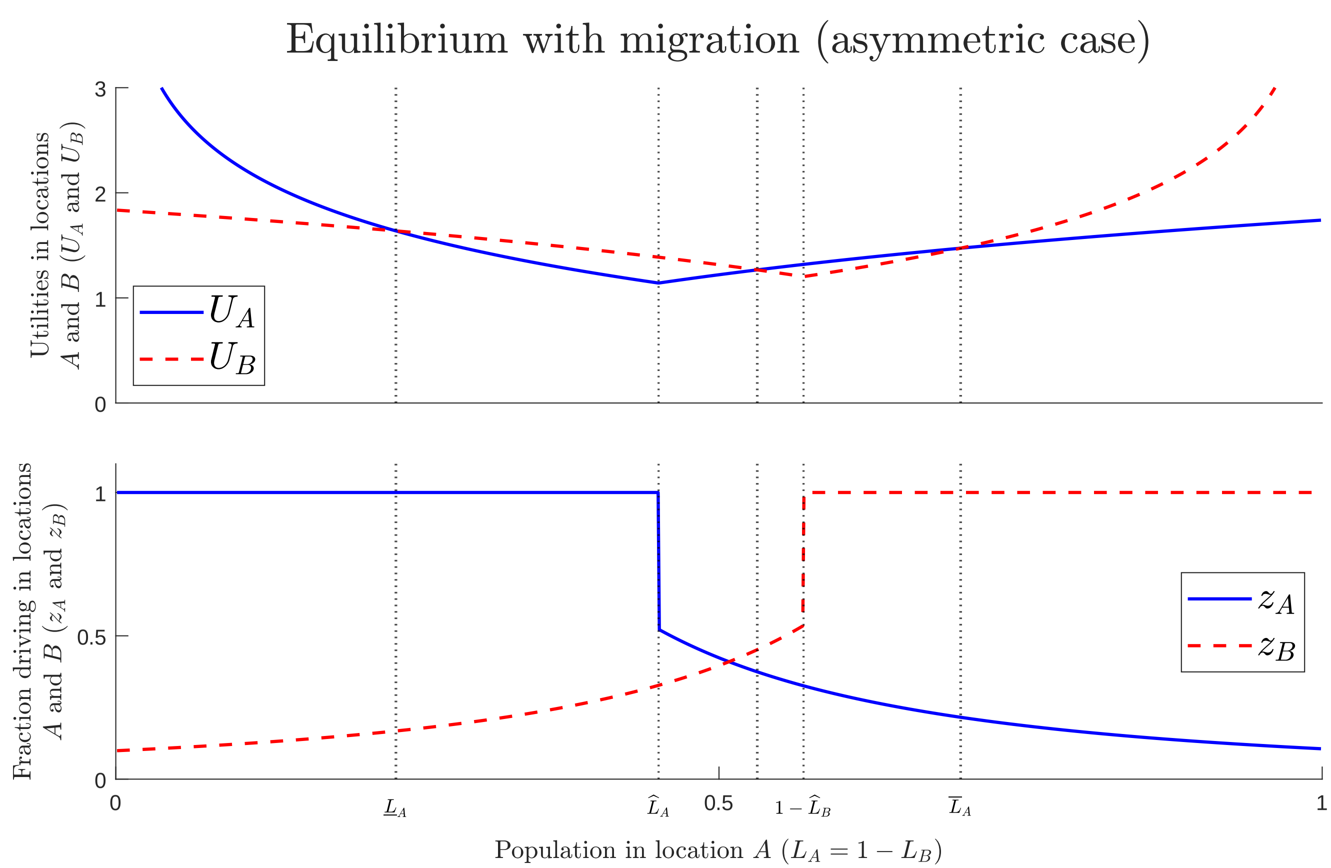

Figure A.4: Equilibrium with migration between two locations, A and B, when $y_B > y_A$.