DISCUSSION PAPER SERIES

DP14717

AUTOMOBILES AND URBAN DENSITY

Francis Ostermeijer, Hans Koster, Jos van Ommeren and Victor Mayland Nielsen

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

ISSN 0265-8003

AUTOMOBILES AND URBAN DENSITY

Francis Ostermeijer, Hans Koster, Jos van Ommeren and Victor Mayland Nielsen

Discussion Paper DP14717
Published 06 May 2020
Submitted 04 May 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

AUTOMOBILES AND URBAN DENSITY

Abstract

How has the rise of the automobile influenced urban areas over the past century? In this paper we investigate the long-run impact of car ownership on urban population density, based on a sample of 232 city observations in 57 countries. Using the presence of a car manufacturer in 1920 as a source of exogenous variation, our IV estimates indicate that car ownership substantially reduces density. A one standard deviation increase in car ownership rates causes a reduction in population density of around 40%. For employment density we find almost identical results. This result has important implications for vehicle taxation, car ownership growth in developing countries, and new transport technologies such as automated vehicles.

JEL Classification: R12, R40

Keywords: Car ownership, vehicle costs, urban density

Francis Ostermeijer - francis.ostermeijer@vu.nl
Vrije Universiteit Amsterdam

Hans Koster - h.koster@vu.nl
Vrije Universiteit Amsterdam and CEPR

Jos van Ommeren - jos.van.ommeren@vu.nl
Vrije Universiteit

Victor Mayland Nielsen - victornielsen@me.com
Significance

Acknowledgements
We thank Jeffrey Kenworthy for generously providing the Millennium Cities Database, Wouter Willemsen for valuable research assistance, Jip Claassens and Maurice de Kleijn for help with GIS, Frederic Robert-Nicoud, Erik Johnson, and Miguel-Angel Garcia-Lopez for comments on earlier drafts, and conference participants at University of Birmingham, UEA Philadelphia and the 2019 SMU conference on Urban and Regional Economics for helpful comments. This work was supported by the Netherlands Organization for Scientific Research (NWO) as part of the 'Spatial and Transport impacts of Automated Driving (STAD)' project 438-15-161.

Automobiles and urban density*

Francis Ostermeijer† Hans R.A. Koster ‡ Jos van Ommeren§
Victor Mayland Nielsen¶

May 6, 2020

Abstract: How has the rise of the automobile influenced urban areas over the past century? In this paper we investigate the long-run impact of car ownership on urban population density, based on a sample of 232 city observations in 57 countries. Using the presence of a car manufacturer in 1920 as a source of exogenous variation, our IV estimates indicate that car ownership substantially reduces density. A one standard deviation increase in car ownership rates causes a reduction in population density of around 40%. For employment density we find almost identical results. This result has important implications for vehicle taxation, car ownership growth in developing countries, and new transport technologies such as automated vehicles.

Keywords: car ownership; vehicle costs; urban density.

JEL codes: R12, R14, R40.

*We thank Jeffrey Kenworthy for generously providing the Millennium Cities Database, Wouter Willemsen for valuable research assistance, Jip Claassens and Maurice de Kleijn for help with GIS, Frédéric Robert-Nicoud, Erik Johnson, and Miguel-Angel Garcia-López for comments on earlier drafts, and conference participants at University of Birmingham, UEA Philadelphia and the 2019 SMU conference on Urban and Regional Economics for helpful comments. This work was supported by the Netherlands Organization for Scientific Research (NWO) as part of the 'Spatial and Transport impacts of Automated Driving (STAD)' project 438-15-161.

†Corresponding author. Department of Spatial Economics, Vrije Universiteit Amsterdam, Netherlands and Tinbergen Institute Amsterdam, Netherlands. Email: francis.ostermeijer@vu.nl.

‡Department of Spatial Economics, Vrije Universiteit Amsterdam, Netherlands and Tinbergen Institute Amsterdam, Netherlands. Research fellow at the National Research University – Higher School of Economics (Russia) and the Tinbergen Institute, Amsterdam, Netherlands. Affiliated to the Centre for Economic Policy Research. Email: h.koster@vu.nl.

§Department of Spatial Economics, Vrije Universiteit Amsterdam, Netherlands and Tinbergen Institute Amsterdam, Netherlands. Research fellow at the Tinbergen Institute, Amsterdam, Netherlands. Email: jos.van.ommeren@vu.nl.

¶Significance, The Hague, Netherlands. Email: victornielsen@me.com.

1 Introduction

"We shall solve the City Problem by leaving the City."

— Henry Ford (1922), *Ford Ideals.*

A defining feature of urbanisation in the 20$^{\text{th}}$ century has been the introduction and rapid, wide-scale adoption of the the automobile. By lowering marginal transport costs and eliminating the need to walk almost entirely, cars allow people to travel longer distances with greater flexibility in terms of routes and schedules. As Henry Ford predicted, this facilitated the decentralisation of cities via the outward expansion of people and firms into the periphery, where land is cheaper, thereby radically changing urban form (Anas et al., 1998; Baum-Snow, 2007; Baum-Snow et al., 2017).

Urban population density is perhaps the most distinguishing characteristic of a city and is a common measure of urban form. Higher density is associated with positive agglomeration economies, public transport efficiency and urban amenities (see Ciccone and Hall, 1996; Glaeser et al., 2001; Rosenthal and Strange, 2004; Glaeser et al., 2008), while lower density linked with higher pollution levels, environmental damage, obesity and segregation of rich and poor (see Brownstone and Golob, 2009; Eid et al., 2008; Zhao and Kaestner, 2010; Couture et al., 2019; Gaigné et al., 2020). Therefore, studying the effect of automobiles on urban density is important in the light that, in most countries, cars are subsidised, implying that car ownership is too high and therefore urban densities may be too low from a welfare perspective (Parry et al., 2007; Au and Henderson, 2006; Brueckner and Helsley, 2011).[1]

In spite of the relevance of this topic, Glaeser and Kahn (2004) argue that we know very little about the long-run effect of car ownership on urban density. This knowledge gap is likely related to the econometric challenge for causal inference of this effect. The first challenge is reverse causality: residents are more likely to use a car in cities with lower urban densities, therefore car ownership rates in these cities may be higher (Bento et al., 2005; Duranton and Turner, 2018; Ewing et al., 2018). Hence, one may overestimate the causal (negative) effect of cars on density if reverse causality is ignored. The second challenge is that urban density is highly persistent over time and is correlated to many difficult-to-observe factors (for example land use planning). So,

[1] For example, in Europe, about 40 percent of all new cars are subsidised through distortionary company car taxation (Van Ommeren and Wentink, 2012), whereas congestion, safety, and environmental externalities are only partially included in the overall price of car use in the US (Parry and Small, 2005).

in order to identify the causal long-run effect of cars on urban density, one requires a *long-term* exogenous shock in car ownership.

We address both challenges using an IV strategy. As an instrument we use the presence of a domestic commercial car manufacturer in 1920, hence when few people owned cars. We provide evidence that countries with a historic car manufacturer currently still pay lower prices for car use and ownership through lower taxation and more roads. Furthermore, we will show that the presence of manufacturer in 1920 is uncorrelated to urban density around that time, which supports our argument that historic car manufacturers are a plausible instrument for car ownership.[2]

Our research design is inspired by Glaeser and Kahn (2004) who were, to our knowledge, the first (and only) to study the causal effect of car ownership on urban density. Using legal origin as an instrument for car ownership, they conclude that cities with lower car ownership rates tend to have higher urban densities. Their main estimate indicates that one additional car per 100 inhabitants is associated with a reduction in urban density of 7.2%. However, as the authors acknowledge themselves, given their limited dataset and identification strategy, these results should be interpreted as suggestive.[3] Our main contribution is to improve on their analysis, by adding new data and introducing a new identification approach. We apply our identification strategy to a cross-section of 232 city observations for 123 cities from 57 countries between 1960 and 2012.

Our work is closely related to a large literature studying the effects of transport infrastructure on the spatial distribution of people and jobs (Baum-Snow, 2007, 2010; Garcia-López et al., 2015; Baum-Snow et al., 2017; Levkovich et al., 2019; Heblich et al., 2020; Gonzalez-Navarro and Turner, 2018). This literature demonstrates that highways are an important driver of decentralisation in the 20$^{\text{th}}$ century, while subways only had a moderate impact. However, highways explain only a portion of car-induced decentralisation.[4] Various other policies, such as vehicles taxes, fuel taxes, and parking regimes, have a strong effect on car ownership and

[2] Up to the extent that one is still concerned that omitted variables bias is an issue, we also estimate fixed-effects models and use the methodology proposed in Oster (2019) to show that our baseline OLS and IV estimates are conservative.

[3] Glaeser and Kahn (2004) employ data from Ingram and Liu (1999), which contains 70 observations for 35 cities in 18 countries (in 1960 and 1980).

[4] Evidence will be provided in Appendix B.4.

use, and thereby on urban density, so estimates of the effect of highways only give a partial view.[5] Hence, we are aim to obtain insight into the overall effect of the automobile, captured by a comprehensive measures such as car ownership, which is the focus of this paper.

The results indicate that one additional car per 100 inhabitants reduces population and employment density by around 2.4% in the long-run. This effect appears to be mainly driven by expansions in the built-up area, and not by population leaving the city, suggesting that cars facilitate low density urban development in the periphery. We use these estimates to gauge the potential effects of growing car ownership rates in developing countries and the introduction of automated vehicles. Applying these estimates, for example, to developing Asian cities indicates that if car ownership increases to similar rates as seen in high-income countries, urban density may fall by over 50% in the long-run. Our estimates are also relevant for high-income countries with low car ownership rates (for example Denmark) as automated vehicles will likely increase access to cars and thereby, in the absence of policy, cause cities to decentralise.

The paper proceeds as follows. In Section 2 we introduce the data and provide some descriptives. In Section 3 we elaborate on the methodology. We report the main results and discuss some implications in Section 4 and finally, Section 5 concludes.

2 Data and descriptives

2.1 Data

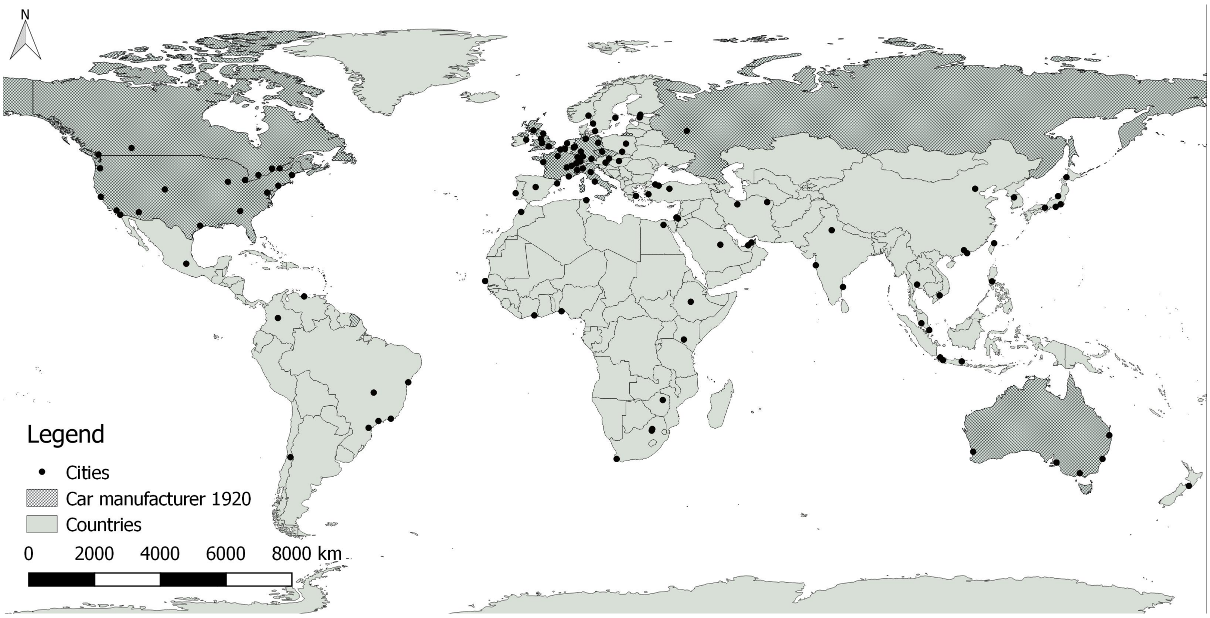
We use several sources of information. The most important source is city level data on population, employment, area size, income and transportation, between 1960 and 2012. We obtain data for 1960 and 1980 from Ingram and Liu (1999), which comprises 69 observations from 35 cities in 18 countries.[6] Data for 1995 comes from the Millennium Cities Database (henceforth MCD1995) and contains information on 100 cities from 51 countries (Kenworthy and Laube, 2001; Kenworthy, 2017).[7] Finally, we also obtain data for 2012 from the Mobility in Cities database (henceforth MCD2012), collected by UITP (2015) using the same methodology as

[5] For example, in Norway there are few highways while car ownership is high.

[6] 25 cities (mainly in developed countries) are collected from Newman and Kenworthy (1989) and are supplemented with 10 other (mostly developing) cities, from various sources. We remove the observation of Guangzhou in 1994 as it overlaps almost exactly with the MCD1995 data.

[7] The dataset contains full information for 89 cities in 41 countries. We impute the missing data points for 11 cities to obtain a complete dataset of 100 cities in 51 countries. In four cities built-up area was missing. In an additional four cities metropolitan GDP was missing and in three cities population density in 1920 was missing. Appendix A.1 documents how we impute these data points.

Figure 1 – Cities in our dataset



MCD1995, which includes 63 cities from 39 countries. A key advantage of these sources is that they rely on a consistent methodology for data collection, which allows us to make accurate comparisons between cities from different countries and time periods. Most importantly, the metropolitan area is consistently defined as the 'commuter belt or labour market region' for all our data and hence captures the bulk of home-work journeys in a city.[8]

Population density is our main measure of urban structure, but we also examine other measures such as employment density and centrality of employment.[9] Population density is measured as the total population in a metropolitan area divided by the total built-up area (in km^2).[10] It therefore captures the density of developed land, accounting for geographical factors such as water and green space which may limit density.

These data sources additionally include information on car ownership per capita, metropolitan GDP per capita, highway length, and the MCD1995 dataset also contains car-related variables such as the average cost of a car trip, annual capital costs of a car, and the number of kilometres of roads and highways. We also collect climate and elevation data from Fick and Hijmans (2017) and Reuter et al. (2007) using longitude and latitude coordinates of each city centroid. This

[8]This is comparable to the OECD (2013) definition of 'functional urban areas', including the hinterland or 'worker catchment area'.

[9]Employment information is only observable for MCD1995 and MCD2012.

[10]Built-up area includes gardens and local parks, urban wasteland, transport infrastructure, recreational, residential, industrial, office, commercial, public utilities, hospitals, schools, cultural areas and sports grounds.

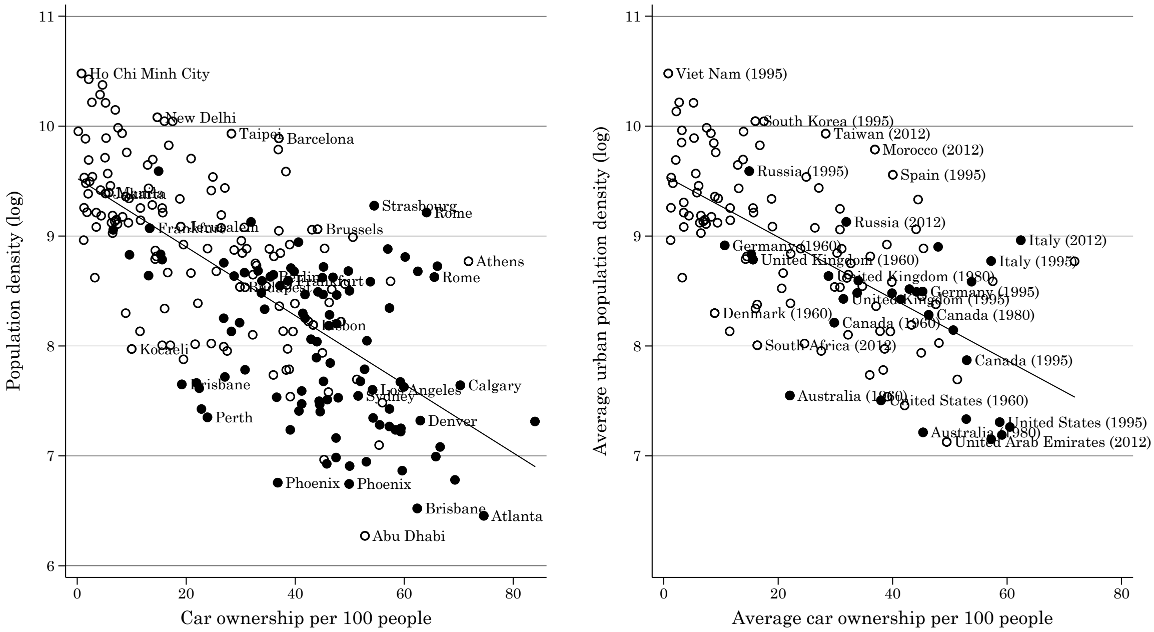
Table 1 – Descriptive statistics

	N	Mean	Std. dev	Min	Max
Population density (pop/km^2)	232	7342.44	6473.72	530.00	35564.53
Employment density (jobs/km^2)	144	3169.18	2712.12	280.00	15127.67
Cars per capita	232	0.32	0.20	0.00	0.84
Car km per capita (100 km)	123	41.90	37.45	0.36	201.97
GDP per capita ($1000s)	232	22.24	21.18	0.22	104.63
Population (millions)	232	4.23	4.92	0.24	37.24
Total built-up area (km^2)	232	874.98	1217.04	44.29	10657.15
Total surface area (km^2)	153	4293.09	8043.50	126.09	57378.00
Built-up to surface area	153	0.36	0.23	0.03	0.93
January temperature (°C)	232	9.31	10.26	-10.40	27.60
July temperature (°C)	232	21.47	5.68	8.80	35.30
Annual precipitation (m)	232	0.97	0.56	0.03	2.93
Altitude (km)	232	0.39	0.53	-0.00	2.60
Ruggedness	232	0.18	0.22	0.00	1.68
Car manufacturer 1920	232	0.37	0.49	0.00	1.00
Country Pop 1913 (millions)	232	49.63	88.85	0.08	437.14
Country GDP per capita 1913 ($1000s)	232	3.77	2.67	0.38	8.38
Pop dens. 1920 (pop/km^2)	232	8689.51	9900.78	1308.93	77736.80
Source: Ingram & Liu 1960	232	0.06	0.24	0.00	1.00
Source: Ingram & Liu 1980	232	0.11	0.32	0.00	1.00
Source: MCD 1995	232	0.53	0.50	0.00	1.00
Source: MCD 2012	232	0.29	0.46	0.00	1.00

allows us to compute average January and July temperatures, annual precipitation, altitude, and ruggedness of the terrain.[11]

To construct our instrument, we collect information on whether a country had a domestic commercial car manufacturer in 1920 by cross-referencing the *Timeline of motor vehicle brands* (Wikipedia, 2018). In Appendix A.3, we document which car manufacturers were present in each specific country, the year they opened and (if relevant) the year they closed down, including primary sources. To complement our historical instrument, we also collect historical country-level data on population and GDP per capita for the year 1913, just before WWI began, from Bolt et al. (2018).[12] Finally, we construct a measure for population density for a representative city in 1920 using data from Goldewijk et al. (2017), as this information is not available using existing sources.[13]

Figure 2 – Population density and car ownership per capita



Notes: The left plot shows city-year data and the right plot shows country-year data. Y-axis is log scaled. Filled (black) circles represent countries with a car manufacturer in 1920. City and country labels are based on minimum, median, and maximum population densities for each bin of 10 cars per 100 people. The solid line represents the bivariate linear regression.

2.2 Descriptive statistics

Descriptive statistics are provided in Table 1. Population density is around 7,300 people per km^2 and the number of cars per capita is 0.32, on average. An average city is large, containing a population of around 4.2 million and spanning a built-up area of around 870 km^2, about 36% of the total city surface area. About one third of cities are located in counties with a car manufacturer in 1920. We provide additional histograms of the main variables of interest in Figure A2 of Appendix A.4. They show, for example, that most cities have population densities below 10,000 people per km^2, with a median of around 5,600.

The cities we focus on are rather large and represent labour market areas. Hence, employment density is highly correlated to population density (the correlation is 0.92). In the analysis, we focus on the effect of car ownership on population density and we repeat the main specifications

[11] We use a similar measure as Burchfield et al. (2006) and calculate ruggedness as the standard deviation of altitude within 50 km from a city's centroid.

[12] Most car manufacturers were present before 1920, therefore it seems appropriate to use information for other historic variables slightly prior to 1920.

[13] We describe the procedure to calculate 1920 population density in Appendix A.2. All data are available upon request.

with employment density and other measures in Section 4.3. Figure 2, left plot, shows that the bivariate relation between car ownership per capita and population density of cities is approximately log-linear and strongly negative. The right plot confirms that message when aggregating our data at the country level.[14] It also shows that countries with historic car manufacturers have notably higher rates of car ownership and lower population densities. In line with common knowledge, US and Australian cities tend to have the highest rates of car ownership and lowest urban densities.

3 Empirical framework

We aim to estimate the long-run causal effect of car ownership per capita on population density. Indexing city i in country j at time t, we set up the following regression equation:

$$\log(D_{ijt}) = \alpha + \beta C_{ijt} + \gamma X_{ijt} + \zeta G_j + \phi_t + \epsilon_{ijt}, \tag{1}$$

where $\log(D_{ijt})$ is the natural logarithm of population density, C_{ijt} represents car ownership per capita, X_{ijt} and G_j are vectors of observed city and country characteristics, ϕ_t are decade fixed effects, and ϵ_{ijt} is an error term. Note that for many cities, we have more than one observation. We address this issue by using weights, with weights inversely proportional to the number of observations per city. As observations do not come from exactly the same year, we control for decade fixed effects, ϕ_t, in all specifications.[15] For all estimates, we cluster standard errors at the country level.

Estimating the marginal effect β with OLS gives consistent estimates of the causal effect of car ownership on urban density, provided that $\text{cov}(C,\epsilon|X,G) = 0$. There are at least two endogeneity concerns when estimating equation (1) by OLS. First, we may omit important variables which affect both population density and car ownership. Second, changes in the urban structure may lead to changes in mobility, resulting in reverse causation as cities with lower densities may induce more car ownership which may in turn cause lower densities (Duranton and Turner, 2018).

To tackle the first issue we include a range of important controls. Higher incomes are correlated

[14] We calculate the urban density in a country, by just considering population in and area size of urban areas in our sample.

[15] This also controls for data source, as each source comes from a different decade.

with higher rates of car ownership (Dargay, 2002) and lower population densities as people demand more space to live in, therefore, we control for the log of GDP per capita at the city level (Margo, 1992).[16] Geographical factors such as the climate, altitude and ruggedness of terrain might also effect car ownership and urban density, for example because larger gardens are more attractive in warmer climates and construction is more expensive in hilly terrain, while active modes of transport like cycling are less likely (Burchfield et al., 2006). Therefore we control for January and July temperatures, precipitation, altitude and ruggedness. The regulatory environment and cultural factors may also play a role in determining attitudes towards car ownership and urban planning (Duranton and Puga, 2015). La Porta et al. (1999, 2008) argue that legal origins influence a broad range of rules and regulations and find that civil law countries tend to be more regulated than common law countries. We therefore control for English, French, German, and Scandinavian legal origins to capture potential correlation between land-use and vehicle regulations which may affect both population density and car ownership.

To examine omitted variable bias of OLS estimates, we also perform a bias-correction approach which allows us to place a bound on the OLS estimate of β, denoted β_{OLS}, in the presence of omitted variables. Oster (2019) shows that a consistent estimate of the bias-adjusted treatment effect can be calculated given assumptions on two key parameters: (i) the proportion of car ownership explained by unobservables relative to observables, δ, and (ii) the maximum variation in the log of population density that can be explained by observables and unobservables, $R^2_{\max}$.[17] To further address concerns related to omitted variable bias, we also estimate fixed-effects models in Section 4.5.1.

In order to tackle the issue of reverse causality, we require a long-term exogenous shock in the use of automobiles. Glaeser and Kahn (2004) apply an IV approach, using legal origin (French civil law) as an instrument for car ownership, so identification is based on country differences.[18]

[16] Note, whether richer households chose to live closer or further from the city centre depends on the sign of the housing and commuting elasticity with respect to income. In European cities, where the urban core is characterized by strong residential and workplace amenities, richer households are likely to locate closer to the centre (Brueckner et al., 1999; Gutiérrez-i Puigarnau et al., 2016).

[17] Oster (2019) recommends to use $R^2_{\max} = 1$ as an upper bound, which implies that any bias will be overstated. Our estimator sometimes delivers multiple solutions if the importance of unobservables is low, *i.e.* $\delta \leq 1$. We then select the solution closest to β_{OLS}, as the alternative solution provides outlier estimates not in line with any other specification. In case that the importance is high, *i.e.* $\delta > 1$, which is the more interesting case, we only get single solutions.

[18] We collect data on legal origin from Appendix B in La Porta et al. (1999).

Legal origin may be argued to be a plausible instrument as it pre-dates the invention of the car and countries with French civil law tend to be more regulated, hence face higher vehicle costs. However, as mentioned above, one criticism is that because countries with French legal origins tended to have more regulation, the instrument may also impact urban density directly via other stricter regulations such as urban planning (La Porta et al., 2008).[19] Another issue is that in our dataset, the instrument appears to be weak (the F-statistic is 2.78).

We therefore propose the presence of a domestic commercial car manufacturer in a country in 1920 as an alternative instrument for car ownership per capita.[20] In the 1920s, few people owned cars. At that time, the US led the world in car manufacturing and ownership. Nevertheless, in the US there were only 8 million registered cars and the ownership rate was only 0.08.[21] At the same time, car manufacturers had substantial political leverage and had a strong lobby, particularly in their home market, to limit vehicle taxes, neglect public transport, and advocate for more road construction and parking in cities (Reich, 1989; Paterson, 2000; Dicken, 2011). After 1920, countries with a historic car manufacturer are therefore likely to have higher rates of car ownership, while the presence of a car manufacturer in 1920 is unlikely to be directly related to urban structure after 1960, other than via car ownership. We will also demonstrate that car manufacturers were not more likely to start up in countries with lower 1920 population density, and that the IV estimates are not sensitive to the inclusion of population in 1920.[22]

We emphasise here that we are not the first study to argue for the impact of the automotive industry on current car policy. For example, Cleff et al. (2005) state that "*[EU] member states having a large car industry tend not to apply a Registration Tax, or they apply a lower Registration tax, while car importing member states tend to levy a higher Registration Tax*".

Probably the most well-known example of how the automotive industry affected car policies

[19]For example, Titman and Twite (2013) find that a countries' legal origin is correlated to the building's lease duration and to the number of high-rise office buildings, and therefore affects urban density directly.

[20]The presence of a car manufacturer at the city level is less likely to be exogenous for two reasons. Firstly, manufacturing plants were large and employed many workers, therefore car manufacturers may have had a direct effect on urban structure at the local level. Secondly, in most countries, national governments determine levels of car and fuel taxation as well as the layout of highways, which are the mechanisms through which car ownership is likely higher.

[21]The US had 8,132,000 registered automobiles and a population of 106,461,000 in 1920 (US Census, 2000). Car ownership was much lower in 1910 and is estimated to be around 500,000. In Section 4.5 we exclude the US as a sensitivity check.

[22]Hence, we address the issue that residents of cities that were more dense before the introduction of the car may have adopted fewer cars and therefore remained more dense.

is the so-called Streetcar Conspiracy, where General Motors and other car manufactures were convicted of monopolising the sale of buses and accused of controlling the transit system in order to dismantle existing streetcar networks, which were replaced by buses (Richmond, 1995). There is also anecdotal evidence that the automotive industry in France launched a powerful lobby in the 1950s against railways in order to promote road construction (Meunier, 2002).

We present further evidence, including several mechanisms, and plausibility of the instrument in Section 4.2.

4 Results

We first present OLS results of the relation between car ownership and population density (Section 4.1), then present evidence on the plausibility of our instrument (Section 4.2) and the IV results (Section 4.3). Finally, we discuss some extensions and perform a range of robustness checks (Section 4.5).

4.1 OLS Results

First we regress the log of population density on cars per capita, controlling only for decade fixed effects. There is a strong and statistically significant negative association. One additional car per 100 inhabitants is associated with a reduction in average population density of around 3%.[23] In column (2) we control for GDP per capita, but the effect of car ownership hardly changes.[24]

In columns (3)-(5), we include historical population density in 1920 and some additional geographical and legal origin controls. The coefficient of historic population density is positive and statistically significant with an elasticity of 0.46, indicating that density is persistent over time. Including 1920 population density reduces the coefficient of interest slightly to −2.4%, meanwhile there is no noticeable effect from the inclusion of climate and terrain controls in column (4).[25] The effect size declines somewhat in column (5) when controls for legal origin are included. The results indicate that countries with French and German legal origins have higher urban densities than countries with English and Scandinavian legal origins. The effect of

[23]This is calculated as $100 \cdot (\exp(\beta) - 1)$.

[24]Note that GDP per capita, or income, does not have a statistically significant effect on population density *when controlling for car ownership*. This result holds regardless of functional form (results from including GDP per capita linearly or with quadratic terms are available upon request). As we will see that income has a strong positive effect on car ownership, in line with Dargay (2002), the overall effect of income on population density appears to be via increased car ownership (see first-stage results in Table 4).

[25]Climate, proxied by summer and winter temperatures, may be correlated to both driving and building types. Both the OLS results and the first-stage results (see Table 4) do not indicate this to be the case.

Table 2 – OLS estimates

	Dep var: Population density (log)				
	(1)	(2)	(3)	(4)	(5)
Cars per 100	-0.0307***	-0.0302***	-0.0243***	-0.0236***	-0.0218***
	(0.00490)	(0.00705)	(0.00541)	(0.00554)	(0.00620)
GDP per capita (log)		-0.0115	0.00803	-0.0373	-0.0346
		(0.0773)	(0.0564)	(0.0593)	(0.0695)
Pop dens. 1920 (log)			0.455***	0.426***	0.356***
			(0.0706)	(0.0705)	(0.0621)
January temperature (°C)				-0.00917	-0.00821
				(0.00665)	(0.00696)
July temperature (°C)				0.0118	0.00869
				(0.0125)	(0.00981)
Annual precipitation (m)				0.182	0.173
				(0.136)	(0.122)
Altitude (km)				-0.182*	-0.185**
				(0.103)	(0.0792)
Ruggedness				0.541**	0.292
				(0.247)	(0.198)
French legal origin					0.476***
					(0.148)
German legal origin					0.302**
					(0.151)
Scandanavian legal origin					-0.274
					(0.170)
Decade FE	Y	Y	Y	Y	Y
R^2	0.502	0.502	0.623	0.660	0.715
No. of countries	57	57	57	57	57
No. of cities	123	123	123	123	123
No. of observations	232	232	232	232	232

Notes: Estimates are weighted by the number of observations per city. Robust standard errors are in parenthesis and are clustered at the country level. Statistical significance is denoted as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

one additional car per 100 inhabitants is associated with a reduction in population density of around 2.2% and is still statistically significant at the 1% level.

Let's now consider Oster's (2019) bias-adjusted estimate. Under the recommended assumption of $\delta = 1$ and $R^2_{\max} = 1$, the bias corrected estimate is -0.043, with a 95% confidence interval between $[-0.060, -0.025]$.[26] We also loosen the assumptions on the δ parameter (see Figure B1 in Appendix B.1). For $\delta \in [0.75,2]$ we calculate the bounds as $[-0.045, -0.038]$ with an average estimated effect of -0.039.[27] This suggests that the OLS coefficient may be somewhat downward biased. We will see that bias-corrected estimates are also somewhat larger than the IV results in Section 4.3.

[26] Standard errors are cluster-bootstrapped (250 replications) based on countries.
[27] As δ increases, the causal estimate converges to around -0.037.

Table 3 – Underlying mechanism

	(1) Car trip cost	(2) Capital cost	(3) Road length	(4) Highway length
Car manufacturer 1920	-0.288*	-0.109	0.377*	0.248
	(0.163)	(0.146)	(0.215)	(0.226)
Controls	Y	Y	Y	Y
Decade FE	Y	Y	Y	Y
R^2	0.613	0.685	0.679	0.581
No. of countries	43	43	50	44
No. of cities	89	92	108	94
No. of observations	89	92	200	124

Notes: Dependent variables are in logs. See Table A3 in Appendix for descriptive statistics. Car trip cost is defined as the direct user cost of an average car trip and includes depreciation, fuel, spare parts, insurance and taxes. Capital cost is defined as the annual fixed costs which includes depreciation, insurance and taxes. Controls are the log of GDP per capita, 1920 population density, January and July temperature, annual precipitation, altitude, ruggedness, and legal origin fixed effects, as in column (5) of Table 2. Road and highway length are per capita. Robust standard errors are in parenthesis and are clustered at the country level. Statistical significance is denoted as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.2 Car manufacturers in 1920

There is a priori no reason why city structure in 1920 was related to the presence of a car manufacturer in 1920 because few people owned cars. Meanwhile, over the subsequent decades, car manufacturers had substantial political leverage and had a strong lobby in their home market to increase car demand (Reich, 1989; Paterson, 2000; Dicken, 2011).

To investigate the latter, we will now examine how the presence of a historic car manufacturer is associated with lower generalised prices for car use at the end of the 20th century.[28] In 1920, nine countries had a domestic commercial car manufacturer.[29] In Table 3 we provide empirical evidence that these countries had lower ownership taxes and lower costs of car use in 1995, even when we control for GDP per capita and other controls.[30] Column (1) indicates that the monetary price of an average car trip, which includes both variable and fixed costs, in 1995 is about 30% lower in countries with historic car manufacturers. Furthermore, these countries have lower annual capital car costs (which include taxes), more roads, and more highway kilometres per capita. However as these effects are imprecisely estimated, they should be interpreted with caution. In Appendix A.3 we also document that in 2005, countries with a historical car manufacturer in Europe charge between 20–50% lower vehicle taxes overall, and face almost zero registration taxes, while fuel taxes are somewhat higher (Kunert and Kuhfeld, 2007).

[28]The MCD 1995 dataset includes information on transport-related costs, however we also have information on road and highway length for a larger sample of cities and time periods.

[29]Australia, Canada, the Czech Republic, France, Germany, Italy, Russia, the United Kingdom, and the United

Table 4 – First-stage results

	Dep var: Cars per 100				
	(1)	(2)	(3)	(4)	(5)
Car manufacturer 1920	27.95***	15.70***	15.94***	15.86***	14.73***
	(4.610)	(3.711)	(3.137)	(2.812)	(3.411)
GDP per capita (log)		9.318***	8.368***	8.244***	8.309***
		(0.747)	(0.729)	(1.059)	(1.081)
Pop dens. 1920 (log)			-4.406**	-4.913**	-4.524**
			(1.891)	(1.871)	(1.889)
Climate controls	N	N	N	Y	Y
Terrain controls	N	N	N	Y	Y
Legal origin FE	N	N	N	N	Y
Decade FE	Y	Y	Y	Y	Y
R^2	0.493	0.726	0.749	0.763	0.765
First-stage F-statistic	36.75	17.89	25.83	31.80	18.64
No. of countries	57	57	57	57	57
No. of cities	123	123	123	123	123
No. of observations	232	232	232	232	232

Notes: Estimates are weighted by the number of observations per city. Robust standard errors are in parenthesis and are clustered at the country level. See Table B1 in Appendix B.6 for table including all controls. Climate controls are; January and July temperatures and annual precipitation, and terrain controls are; altitude and ruggedness as in column (5) of Table 2. Statistical significance is denoted as: * p < 0.1, ** p < 0.05, *** p < 0.01. Kleibergen-Paap F-statistic is presented.

We present the first-stage results of the IV estimates in Table 4. Columns (1)-(5) indicate that the presence of a commercial car manufacturer in 1920 is a strong instrument.[31] The Kleibergen-Paap first-stage F-statistic is 18.58 in the last (and preferred) specification. The instrument has the expected positive sign: countries with a commercial car manufacturer owned around 15 more cars per capita (or around 50% more than the mean city in our sample).

We come now back to our claim that a priori, there is no reason why population density in 1920 was related to the presence of a car manufacturer in 1920. In Table 5 we examine whether car manufacturers in 1920 were more likely to be present in cities with lower population densities, as arguably, the instrument is more convincing if it is not related to population density in 1920. Columns (1) and (2) indicate that the presence of a historic car manufacturer is not correlated to historic population density, independent of whether we include controls. If anything, the point estimate in the preferred specification in column (2) suggests that population density was higher in countries with a car manufacturer in 1920.[32]

States. See Appendix A.3.

[30] We have fewer observations here than in the main analysis because of missing information.

[31] Note that car ownership at the city level strongly increases with GDP per capita, and falls with historic population density).

[32] This is not surprising given that, for example, US cities faced higher urban densities until 1950, after which they began to decline (Kim, 2007).

Table 5 – Additional validity tests of instrument

	Population density 1920 (log)		Population density (log)	
	(1)	(2)	(3)	(4)
Car manufacturer 1920	-0.227	0.249	-0.819***	-0.363**
	(0.270)	(0.206)	(0.289)	(0.158)
Decade FE	Y	Y	Y	Y
Controls	N	Y	N	Y
R^2	0.0520	0.281	0.232	0.661
No. of countries	57	57	57	57
No. of cities	123	123	123	123
No. of observations	232	232	232	232

Notes: Estimates are weighted by the number of observations per city. The dependent variable is historic (1920) or observed (1960-2012) population density in logs. Robust standard errors are in parenthesis and are clustered at the country level. Controls are the log of GDP per capita, January and July temperature, annual precipitation, altitude, ruggedness, legal origin fixed effects, and in columns (3) and (4) we also include 1920 population density, as in column (5) of Table 2. Statistical significance is denoted as: * p < 0.1, ** p < 0.05, *** p < 0.01.

We also estimate the reduced-form effect on current population density (after 1960) in columns (3) and (4). The results indicate that the presence of a car manufacturer in 1920 has a strong, negative, and statistically significant, effect on population density many years later.[33] Hence, historic car manufacturers are not related to historic population density, but have a causal effect on current density. In the following 2SLS analysis we control for historic population density, but we emphasise that the results of Table 5 imply that our instrument is also unconditionally valid, *i.e.* when not controlling for historic population density, we will get the same effect for car ownership.

4.3 IV Results

In Table 6 we provide the 2SLS results using the presence of a historic car manufacturer as an instrument. The coefficient of interest remains quite stable to the inclusion of control variables and are of a similar order of magnitude to the OLS estimates. Furthermore, the control variables have a similar effect as compared to the OLS estimates in Table 2.

The preferred specification in column (5) indicates that one additional car per 100 inhabitants is associated with a reduction in population density of around 2.4%. A Hausman test does not reject the null hypothesis that the IV and OLS coefficient on cars per 100 in column (5) is significantly different from each other.[34] The IV estimate in column (5) is in between the

[33]A formal test for the difference between the effect of car manufacturers in (2) and (4) rejects the null hypothesis that there is no difference at the 95% confidence level. The difference is -0.62, with a standard error of 0.26 and a corresponding t-statistic of -2.40.

[34]We perform a cluster-robust Hausman test with 250 bootstrap replications (Cameron and Trivedi, 2005). The test statistic is $\chi^2(1) = 0.02$, with a corresponding p-value of 0.89.

Table 6 – 2SLS estimates

	Dep var: Population density (log)				
	(1)	(2)	(3)	(4)	(5)
Cars per 100	-0.0293***	-0.0268**	-0.0283***	-0.0261***	-0.0247**
	(0.00718)	(0.0133)	(0.00989)	(0.0100)	(0.0102)
GDP per capita (log)		-0.0542	0.0550	-0.0103	-0.00380
		(0.146)	(0.110)	(0.109)	(0.114)
Pop dens. 1920 (log)			0.439***	0.414***	0.346***
			(0.0726)	(0.0752)	(0.0693)
Climate controls	N	N	N	Y	Y
Terrain controls	N	N	N	Y	Y
Legal origin FE	N	N	N	N	Y
Decade FE	Y	Y	Y	Y	Y
First-stage F-statistic	36.75	17.89	25.83	31.80	18.64
No. of countries	57	57	57	57	57
No. of cities	123	123	123	123	123
No. of observations	232	232	232	232	232

Notes: Estimates are weighted by the number of observations per city. See Table B2 in Appendix B.6 for table including all controls. Climate controls are; January and July temperatures and annual precipitation, and terrain controls are; altitude and ruggedness as in column (5) of Table 2. Robust standard errors are in parenthesis and are clustered at the country level. Kleibergen-Paap F-statistic is presented. Statistical significance is denoted as: * p < 0.1, ** p < 0.05, *** p < 0.01.

OLS estimate in Table 2 and the bias-corrected estimate using Oster's (2019) method in Section 4.1. Apparently, the reverse causality issue is too small to have consequences for the estimates, at least at the city level.[35] The estimate implies that a one standard deviation increase in car ownership (20 cars per 100 inhabitants) is associated with a reduction in population density of around 40%.[36] We find a smaller effect than in Glaeser and Kahn (2004) who find an effect size around three times as large.[37]

4.4 Extensions

4.4.1 Other dependent variables

In Table 7 we present results separating population density into population and area size of the city, and consider two additional dependent variables: employment density and employment centrality, defined by the number of jobs in the CBD. Columns (1) and (2) indicate that although the effect of car ownership rates is not statistically significant when regressing both variables separately, the point estimate of cars per 100 on the log of built-up area is 0.025, which is essentially identical to the estimate of cars on population density in column (5) of Table 6, meanwhile the point estimate on population is essentially zero. This is in line with the sprawl

[35] Duranton and Turner (2018) find that urban density has a small negative effect on vehicle kilometres driven, however their study is at the household level rather than the city level.

[36] This can be calculated as: $\exp(\hat{\beta} \times \Delta C) - 1 = exp(-0.0245 \times 20) - 1 = -0.387$ or 39%.

[37] Here we refer to their coefficient in Table 6, column (3), which is -0.075 or 7.2%.

Table 7 – 2SLS Sensitivity checks: Other dependent variables

	(1) Population	(2) Area	(3) Emp. density	(4) Prop. Jobs CBD
Cars per 100	0.000371	0.0250	-0.0216*	-0.0112
	(0.0158)	(0.0212)	(0.0111)	(0.0124)
Controls	Y	Y	Y	Y
Decade FE	Y	Y	Y	Y
First-stage F-statistic	18.64	18.64	14.53	17.54
No. of countries	57	57	53	46
No. of cities	123	123	112	112
No. of observations	232	232	144	93

Notes: Dependent variables are in logs. Estimates are weighted by the number of observations per city. Controls are the log of GDP per capita, 1920 population density, January and July temperature, annual precipitation, altitude, ruggedness, and legal origin fixed effects, as in column (5) of Table 2. Robust standard errors are in parenthesis and are clustered at the country level. Kleibergen-Paap F-statistic is presented. Statistical significance is denoted as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

hypothesis as it suggests that the overall effect of cars on population density appears to be via cars causing cities to spread out further (Glaeser and Kahn, 2004; Nechyba and Walsh, 2004; Su and DeSalvo, 2008).

As mentioned in Section 2, population and employment density are highly correlated – therefore one expects similar effects for population and employment density.[38] Column (3) indicates that one additional car per 100 inhabitants causes a reduction in employment density of around 2.1%, but this effect is just not statistically significant at the 5% level, and only significant at the 10% level. We also examine whether car ownership rates reduce the proportion of jobs in the CBD in column (4), however we do not find a statistically significant effect.

4.4.2 Heterogeneous effects

We also investigate whether the effect of car ownership on population density varies by *(i)* infrastructure quality, *(ii)* level of GDP, and *(iii)* origin of law system. Table 8 reports the results.[39] Columns (1) and (2) suggest that roads and highways appear to have a complementary effect, so both roads *and* cars are important in facilitating urban decentralisation. Interestingly, roads also have a statistically significant and negative effect on population density, while the effect of highways is a lot smaller. This may be because road space occupies a large share of land, which can no longer be used for residential purposes.

In column (3) we find some evidence that the effect of car ownership on population density is

[38] Baum-Snow (2010) also finds similar effects of highways on the decentralisation of firms and households.

[39] To improve efficiency of the estimates, we include interactions of the fitted values from the first-stage with road length, GDP per capita and legal origin, see Levkovich et al. (2019) which compares different methodologies.

Table 8 – 2SLS sensitivity checks: Heterogeneity

	Dep var: Population density (log)			
	(1) Road length	(2) Highway length	(3) GDP	(4) Legal origin
Cars per 100 (demeaned)	-0.0178**	-0.0146	-0.0176	-0.0330***
	(0.00859)	(0.0106)	(0.0113)	(0.00921)
× Road length demeaned (log)	-0.00781***			
	(0.00175)			
× Highway length demeaned (log)		-0.00772***		
		(0.00235)		
× GDP per capita demeaned (log)			-0.00438**	
			(0.00215)	
× French origin				0.0226***
				(0.00449)
× German origin				0.00218
				(0.00675)
× Scandanavian origin				-0.0111
				(0.0131)
GDP per capita demeaned (log)			-0.140	
			(0.149)	
Road length demeaned (log)	-0.111***			
	(0.0397)			
Highway length demeaned (log)		-0.0460		
		(0.0553)		
GDP per capita (log)	-0.118	-0.176		-0.00991
	(0.118)	(0.150)		(0.108)
French legal origin	0.333**	0.475***	0.477***	0.493***
	(0.156)	(0.174)	(0.153)	(0.131)
German legal origin	0.123	0.235	0.316	0.229
	(0.204)	(0.211)	(0.200)	(0.185)
Scandanavian legal origin	-0.265	-0.299	-0.178	-0.264
	(0.227)	(0.308)	(0.266)	(0.219)
Controls	Y	Y	Y	Y
Decade FE	Y	Y	Y	Y
First-stage F-statistic	18.64	18.64	18.64	18.64
No. of countries	50	44	57	57
No. of cities	200	124	232	232

Notes: Estimates are weighted by the number of observations per city. Controls are 1920 population density, January and July temperature, annual precipitation, altitude, ruggedness, and legal origin fixed effects, as in column (5) of Table 2. Robust standard errors are in parenthesis and are clustered at the country level. Kleibergen-Paap F-statistic is presented. Statistical significance is denoted as: * p < 0.1, ** p < 0.05, *** p < 0.01.

larger in countries with a higher GDP. This result might be reasonable as richer countries are likely to invest in road infrastructure that strengthen the effect of car ownership.

Finally, column (4) indicates that while the effect of car ownership appears to be similar for countries with English, German, and Scandinavian legal origins, countries with French legal origins appear to have significantly smaller effects (around half). This indicates that countries with French legal origins are not only more regulated in terms of taxation (Glaeser and Kahn,

Table 9 – Additional sensitivity checks: fixed effects models

	Dep var: Population density (log)				
	(1)	(2)	(3)	(4)	(5)
	IV	OLS	OLS	OLS	OLS
Cars per 100	-0.0303	-0.0198**	-0.0145*	-0.00374	-0.0156***
	(0.0309)	(0.00840)	(0.00811)	(0.00434)	(0.00488)
Controls	Y	Y	Y	Y	Y
Decade FE	Y	Y	Y	Y	Y
Area FE	Continent (6)	Continent (6)	Country	City	City
R^2		0.742	0.888	0.964	0.967
First-stage F-statistic	3.895				
No. of countries	57	57	57	34	15
No. of cities	107	107	107	60	29
No. of observations	232	232	232	169	77

Notes: Estimates are weighted by the number of observations per city. Controls are the log of GDP per capita, 1920 population density, January and July temperature, annual precipitation, altitude, ruggedness, and legal origin fixed effects, as in column (5) of Table 2. Robust standard errors are in parenthesis and are clustered at the country level. Kleibergen-Paap F-statistic is presented. Statistical significance is denoted as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

2004) but also have more planning restrictions.

4.5 Robustness

The IV results indicate that one additional car per 100 inhabitants reduces population density at the city level by 2.4% in the long-run. In this subsection we perform a wide range of robustness checks, and provide some tentative evidence on the middle-run effect.

4.5.1 Fixed-effects models

Up to now, we have exploited (mainly) cross-sectional variation in population density and car ownership. We assess the sensitivity of our results to various alternative types of variation, by gradually including a more detailed set of fixed effects. The first column of Table 9 shows IV results and columns (2)-(5) show OLS results. In column (1) and (2) we include continent fixed effects. The limitation of this analysis is that because our instrument varies at the country level, the degrees of freedom at the country level is strongly reduced. As a consequence, the first-stage F-statistic falls below 10 and the coefficient of interest becomes imprecise. Nevertheless, including continent fixed effects leads to similar (slightly larger) results than our main specification. The OLS regression with continent fixed effects in column (2) delivers a very similar estimate as compared to the baseline OLS estimate.

Next we include country fixed effects, therefore we only exploit cross-sectional variation *within* a country. This has the advantage that we can control for unobserved factors at the national level such as regulations that effect both vehicle ownership and urban density, however as our instrument does not vary at the city level, we are unable to correct for reverse causality. So far, our results have shown that the OLS and IV results are remarkably similar, and the OLS results are generally more conservative, therefore we tentatively perform this analysis to check the robustness of our results, but urge caution when interpreting the coefficients as causal effects. The results in column (3) indicate that the effect is roughly similar, one additional car is associated with a reduction in population density of around -1.5% and is statistically significant at the 10% level.

We then exploit temporal variation by including city fixed effects in column (4). This provides a tentative estimate of the middle-run effects. The estimate however becomes close to zero and is not statistically significant. An issue with this specification is that urban density changes slowly and information about cities that are observed with few years in between comes from different sources, implying substantial measurement error. The latter is problematic, because the downward bias due to measurement error is compounded with panel data. In order to overcome the inconsistency, Cameron and Trivedi (2005) recommend using longer differences, therefore in specification (5) we only select observations for cities that are at least 20 years apart. This leaves us with 77 cities and 29 observations. The coefficient is -1.5% and statistically significant at the 1% level. This is smaller in magnitude than our main IV result, but around the same size as the OLS result when including country fixed effects, suggesting that the middle-run effect may be about half of the size of the long-run effect.

4.5.2 Other robustness checks

In Appendix B.3 we consider alternative proxies for automobile use. We include motorbike use as a separate variable and together with car ownership, leading to similar results. We also consider measuring car usage directly, by using car km per capita and car km per car. For car km per capita we find almost identical results. However, for car kilometres per car the results are imprecise because of a weak first stage. This may be because car manufacturers focus lobbying efforts on car ownership and purchase, while having smaller effects of car usage such as fuel taxes, which is corroborated by our analysis on European countries in Section 4.2. Still, despite weak instruments, we find a negative second-stage coefficient of car use on population density of the same order of magnitude as the baseline estimates.

A large literature has investigated the effects of highways on population density and decentralisation. In Appendix B.4 we further investigate whether the effect of car ownership and infrastructure on population density are complementary. However, note that roads and highways cause cities to decentralise only when there is sufficient car use. We show that roads and highways have the expected negative effect on population density. The effect of car ownership is only reduced by about 20%, suggesting that we also expect decentralisation to happen when the infrastructure is still immature (say in cities in Sub-Saharan Africa).

Our data is composed of three main data sources, obtained from Ingram and Liu (1999), Kenworthy and Laube (2001), and UITP (2015). In Appendix B.5 we assess the sensitivity of our results to the various data sources. There are two important observations from this exercise. The first-stage F-statistic becomes weaker for more recent data; and the second-stage estimate becomes lower. This suggests that car manufacturers were more powerful in the 1960s and 1970s in influencing policy, which is in line with anecdotal evidence. The reduction in the magnitude of the second-stage coefficient appears to be driven by the inclusion of other cities and countries. This suggests that the sample of cities in Ingram and Liu (1999) is not completely representative.

We consider various other robustness checks in Appendix B.6. We first make sure that the effect of car manufacturers in 1920 on car ownership is not confounded by GDP per capita around that time, as car manufacturers in 1920 tended to be present in higher income countries. Although the coefficient of interest becomes somewhat imprecise, the magnitude of the point estimate increases slightly and is very close to our preferred baseline estimate. We also investigate whether the results are robust to the inclusion of a quadratic polynomial of current GDP per capita, but this does not materially influence the results.

One may be worried that our results are driven by cities in the US due to their firmly established car culture and the abundance of land for urban expansion. However, when we exclude the 12 cities in the US from our sample (27 observations) the effect size decreases only slightly to -2%.

As an instrument we use a dummy whether there is a car manufacturer in 1920. Alternatively, we also consider to use a dummy whether there is a car manufacturer in 1910 or 1930. This leads to similar results, although the results are statistically stronger once we use the dummy indicating whether there is a car manufacturer in 1930.

4.6 Implications

Overall, our preferred estimate from column (5) in Table 6 implies that an increase in car ownership of one car per 100 inhabitants leads to a reduction in population density of around 2.4%. In this section, we apply this estimate to gauge the potential effects of growing car ownership rates in developing countries and the introduction of autonomous vehicles on urban density.

4.6.1 Growing car ownership in developing countries

In 1995, cities located in developing Asian countries owned substantially fewer cars per capita and faced higher population densities (see Table A4 in Appendix). Applying our estimates suggests that if car ownership increases to similar rates as seen in western Europe, urban density would fall by around 50% in the long-run, while if car ownership rates reach levels seen in North America and Oceania, density would even fall by around 60%.

We have applied our estimates for three Chinese cities in our dataset: Beijing, Guangzhou, and Shanghai. In 1995 average car ownership in these cities was 2.6 cars per 100 inhabitants which grew to 17.5 by 2010.[40] According to our results this would result in a reduction in population density of around 30%, whereas the actual reduction was around 60% (it declined from 14,600 to 5,600 people per km^2). This suggests that changes in car ownership explain about half the reduction in population density.[41]

4.6.2 Automated vehicles

Our estimates are also relevant in the broader context of future transport developments such as fully automated vehicles (AVs) which are expected to increase access to cars. These results are particularly relevant to cities with relatively high incomes, but low levels of car ownership such as Copenhagen (Denmark) and Tokyo (Japan). Currently, car use is limited by ownership, however, AVs are expected to reduce the fixed costs of owning a car and thereby may substantially increase vehicle access. In the absence of policy, our results suggest that cities are expected to become more decentralised.

Fagnant and Kockelman (2015) assume that AVs are expected to increase vehicle kilometres travelled (VKT) by $10 - 20\%$ in the US. We use our estimate for the change in VKT per capita

[40] Information for the year 2010 is gathered from ITF (2017) and Demographia (2010).

[41] As the period between 1995 and 2010 is relatively short, it is plausible that the contribution of car ownership is less.

in column (4) of Table B3 in Appendix B.3 and consider these changes as lower and upper bounds of the effects of AVs. In scenario (A) effective car ownership for an average city in our dataset increases by 10%, or 420 km per person, leading to a decline in population density of around 6.5%. In scenario (B) we consider a more extreme situation where equivalent car use increases by 20%, which is expected to result in population density declining by around 12.5% in the long-run.

While these estimates provide a rough indication of the potential effects of AVs, there may be reasons to expect that they may be over- or under-estimates. On the one hand, because AVs can be shared and therefore do not require car ownership, this this may free up vast amounts of parking space in inner cities which could be used for residential and other purposes, while on the other hand, because commuters can engage in other activities in the vehicle, such as sleeping or working, this might lead to longer commutes than currently tolerated (Pudāne et al., 2019).

5 Conclusions

Cars have dominated the urban landscape over the past century. In this paper we investigate the long-run impact of car ownership on urban form, in particular on population density, in an international sample of cities. Using the presence of a car manufacturer in 1920 as a source of exogenous long-term variation in vehicle costs, our IV estimates indicate that higher car ownership rates, induced via lower ownership costs, substantially reduce densities. A one standard deviation increase in car ownership rates (or 20 cars per 100 inhabitants) causes a reduction in density of around 40% in the long-run. Disentangling this effect between population and city size suggests that the major driver of this reduction in urban density is via the outward expansion of the city as the size of urban areas increase. Furthermore, we find that the effects are larger in cities with more roads and highways, and a higher income; they are lower in countries with French legal origins, which may have stricter vehicle taxation and land use regulations.

Our findings suggest that unpriced market failures in the car market have additional spillovers on urban density. This has implications for the key benefits of living and working in a city, and may justify higher taxes on private vehicle ownership and use in order to increase the benefits associated with higher densities, such as positive agglomeration economies and public transport efficiency, and decrease the costs associated with lower densities, such as pollution and environmental damage. Furthermore, the paper also has implications for expected urban growth in developing countries, where car ownership rates and populations are rapidly increasing, and future transport technologies such as automated vehicles, which are expected to dramatically reduce the costs of using a private vehicle.

References

Anas, A., Arnott, R., and Small, K. A. (1998). Urban spatial structure. *Journal of Economic Literature*, 36(3):1426–1464.

Angel, S., Parent, J., Civco, D. L., and Blei, A. M. (2010). *Persistent Decline in Urban Densities: Global and Historical Evidence of 'Sprawl'*. Lincoln Institute of Land Policy.

Au, C.-C. and Henderson, J. V. (2006). Are chinese cities too small? *The Review of Economic Studies*, 73(3):549–576.

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.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. (2017). Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics*, 99(3):435–448.

Bento, A. M., Cropper, M. L., Mobarak, A. M., and Vinha, K. (2005). The effects of Urban spatial structure on travel demand in the United States. *Review of Economics and Statistics*, 87(3):466–478.

Bolt, J., Inklaar, R., de Jong, H., and van Zanden, J. (2018). Maddison project database, version 2018. *Rebasing 'Maddison': new income comparisons and the shape of long-run economic development.*

Brownstone, D. and Golob, T. F. (2009). The impact of residential density on vehicle usage and energy consumption. *Journal of Urban Economics*, 65(1):91–98.

Brueckner, J. K. and Helsley, R. W. (2011). Sprawl and blight. *Journal of Urban Economics*, 69(2):205–213.

Brueckner, J. K., Thisse, J.-F., and Zenou, Y. (1999). Why is central paris rich and downtown detroit poor?: An amenity-based theory. *European economic review*, 43(1):91–107.

Burchfield, M., Overman, H. G., Puga, D., and Turner, M. A. (2006). Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics*, 121(2):587–633.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: methods and applications.* Cambridge university press.

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

Cleff, T., Heneric, O., and Spielkamp, A. (2005). Regulation and industrial policy. In Heneric, O., Licht, G., and Sofka, W., editors, *Europe's automotive industry on the move: competitiveness in a changing world*, pages 157–190. Springer-Verlag, Mannheim.

Couture, V., Gaubert, C., Handbury, J., and Hurst, E. (2019). Income growth and the distributional effects of urban spatial sorting. *NBER Working Paper No. 26142.*

Dargay, J. M. (2002). Determinants of car ownership in rural and urban areas: a pseudo-panel analysis. *Transportation Research Part E: Logistics and Transportation Review*, 38(5):351–366.

Demographia (2010). Demographia World Urban Areas: Population Projections, UN. Technical report.

Dicken, P. (2011). *Global shift: Mapping the changing contours of the world economy.* The Guilford Press, 6 edition.

Dobson, J. E., Bright, E. A., Coleman, P. R., Durfee, R. C., and Worley, B. A. (2000). Landscan: a global population database for estimating populations at risk. *Photogrammetric engineering and remote sensing*, 66(7):849–857.

Duranton, G. and Puga, D. (2015). Urban land use. In *Handbook of regional and urban economics*, volume 5, pages 467–560. Elsevier.

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

Eid, J., Overman, H. G., Puga, D., and Turner, M. A. (2008). Fat city: Questioning the relationship between urban sprawl and obesity. *Journal of Urban Economics,*, 63(2):385–404.

Ewing, R., Hamidi, S., Tian, G., Proffitt, D., Tonin, S., and Fregolent, L. (2018). Testing Newman and Kenworthy's theory of density and automobile dependence. *Journal of Planning Education and Research*, 38(2):167–182.

Fagnant, D. J. and Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, 77:167–181.

Fick, S. E. and Hijmans, R. J. (2017). Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12):4302–4315.

Henry Ford (1922). *Ford Ideals: Being a Selection from" Mr. Ford's Page" in The Dearborn Independent.* Dearborn Publishing Company.

Gaigné, C., Koster, H. R. A., Moizeau, F., and Thisse, J. (2020). Who lives where in the city? Amenities, commuting and income sorting. *CEPR Discussion Paper 11958.*

Garcia-López, M.-À., Holl, A., and Viladecans-Marsal, E. (2015). Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities. *Journal of Urban Economics*, 85:52–67.

Glaeser, E. L. and Kahn, M. E. (2004). Sprawl and urban growth. In *Handbook of Regional and Urban Economics*, volume 4, pages 2481–2527. Elsevier.

Glaeser, E. L., Kahn, M. E., and Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics*, 63(1):1–24.

Glaeser, E. L., Kolko, J., and Saiz, A. (2001). Consumer city. *Journal of Economic Geography*, 1(1):27–50.

Goldewijk, K. K., Beusen, A., Doelman, J., and Stehfest, E. (2017). Anthropogenic land use estimates for the holocene–hyde 3.2. *Earth System Science Data*, 9(1):927–953.

Gonzalez-Navarro, M. and Turner, M. A. (2018). Subways and urban growth: Evidence from earth. *Journal of Urban Economics*, 108:85–106.

Gutiérrez-i Puigarnau, E., Mulalic, I., and van Ommeren, J. N. (2016). Do rich households live farther away from their workplaces? *Journal of Economic Geography*, 16(1):177–201.

Heblich, S., Redding, S. J., and Sturm, D. M. (2020). The making of the modern metropolis: Evidence from London. *Quarterly Journal of Economics*.

Ingram, G. K. and Liu, Z. (1999). *Motorization and the Provision of Roads in Countries and Cities*. The World Bank.

ITF (2017). International Transport Forum Transport Outlook 2017, OECD.

Kenworthy, J. and Laube, F. (2001). *The Millenium Cities Database for Sustainable Transport*. International Union (Association) of Public Transport (UITP) and the Institute for Sustainability and Technology Policy (ISTP).

Kenworthy, J. R. (2017). The good, the bad and the ugly in urban transport: Comparing global cities for dependence on the automobile. *Methods for Sustainability Research*, pages 46–62.

Kim, S. (2007). Changes in the nature of urban spatial structure in the United States, 1890–2000. *Journal of Regional Science*, 47(2):273–287.

Kunert, U. and Kuhfeld, H. (2007). The diverse structures of passenger car taxation in europe and the eu commissions proposal for reform. *Transport Policy*, 14(4):306–316.

La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2008). The economic consequences of legal origins. *Journal of Economic Literature*, 46(2):285–332.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. (1999). The quality of government. *The Journal of Law, Economics, and Organization*, 15(1):222–279.

Lahmeyer, J. (2006). Population statistics: historical demography of all countries, their divisions and towns. http://populstat.info/. Accessed: 28-09-2018.

Levkovich, O., Rouwendal, J., and van Ommeren, J. (2019). The impact of highways on population redistribution: the role of land development restrictions. *Journal of Economic Geography*. lbz003.

Margo, R. A. (1992). Explaining the postwar suburbanization of population in the United States: The role of income. *Journal of Urban Economics*, 31(3):301–310.

Meunier, J. (2002). *On the fast track: French railway modernization and the origins of the TGV, 1944-1983*. Greenwood Publishing Group, Santa Barbara CA.

Nechyba, T. J. and Walsh, R. P. (2004). Urban sprawl. *Journal of Economic Perspectives*, 18(4):177–200.

Newman, P. G. and Kenworthy, J. R. (1989). *Cities and automobile dependence: An international sourcebook*. Aldershot, England: Gower Publishing.

OECD (2013). Definition of Functional Urban Areas (FUA) for the OECD Metropolitan Database.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

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., Walls, M., and Harrington, W. (2007). Automobile externalities and policies. *Journal of economic literature*, 45(2):373–399.

Paterson, M. (2000). Car culture and global environmental politics. *Review of International Studies*, 26(2):253–270.

Patterson, T. and Kelso, N. V. (2012). World urban areas, landscan, 1:10 million (2012). https://geo.nyu.edu/catalog/stanford-yk247bg4748. Accessed: 11-12-2018.

Pudāne, B., Rataj, M., Molin, E. J., Mouter, N., van Cranenburgh, S., and Chorus, C. G. (2019). How will automated vehicles shape users' daily activities? Insights from focus groups with commuters in the Netherlands. *Transportation Research Part D: Transport and Environment*, 71:222–235.

Reich, S. (1989). Roads to follow: regulating direct foreign investment. *International Organization*, 43(4):543–584.

Reuter, H. I., Nelson, A., and Jarvis, A. (2007). An evaluation of void-filling interpolation methods for srtm data. *International Journal of Geographical Information Science*, 21(9):983–1008.

Richmond, J. (1995). *Transport of delight: the mythical conception of rail transit in Los Angeles*. University of Akron Press, Akron OH.

Rosenthal, S. S. and Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In *Handbook of Regional and Urban Economics*, volume 4, pages 2119–2171. Elsevier.

Schneider, A., Friedl, M. A., McIver, D. K., and Woodcock, C. E. (2003). Mapping urban areas by fusing multiple sources of coarse resolution remotely sensed data. *Photogrammetric Engineering & Remote Sensing*, 69(12):1377–1386.

Smith, D. A. (2017). Visualising world population density as an interactive multi-scale map using the global human settlement population layer. *Journal of Maps*, 13(1):117–123.

Su, Q. and DeSalvo, J. S. (2008). The effect of transportation subsidies on urban sprawl. *Journal of Regional Science*, 48(3):567–594.

Titman, S. and Twite, G. (2013). Urban density, law and the duration of real estate leases. *Journal of Urban Economics*, 74:99–112.

UITP (2015). Mobility in Cities Database 2015. International Association of Public Transport.

US Census (2000). 2. Population and 1027. Motor Vehicle Registrations. https://www.allcountries.org/uscensus. Accessed: 07-03-2019.

Van Ommeren, J. and Wentink, D. (2012). The (hidden) cost of employer parking policies. *International Economic*

Review, 53(3):965–978.

Wikipedia (2018). Timeline of motor vehicle brands. https://en.wikipedia.org/wiki/Timeline_of_motor_vehicle_brands. Accessed: 21-08-2018.

Zhao, Z. and Kaestner, R. (2010). Effects of urban sprawl on obesity. *Journal of Health Economics*, 29(6):779–787.

Appendix A Data appendix

A.1 Appending data sources

Our main datasources are from Ingram and Liu (1999), Kenworthy and Laube (2001), and UITP (2015). Several important aspects should be noted. As some data points were missing in the original data, we imputed these observations using the most reliable data available.

A.1.1 Population density missing

The MCD 2012 dataset (UITP, 2015) was missing population density information for Addis Ababa (Ethiopia), Dublin (Ireland), Glasgow (Scotland), Izmir (Turkey), Johannesburg and Tshwane (South Africa), Mumbai (India), Nairobi (Kenya), Tehran (Iran), and Tokyo (Japan). Therefore, we collect data on the average population density of the built-up area in the closest available year (2015) from Smith (2017).[42]

Additionally, the MCD 1995 data (Kenworthy and Laube, 2001) did not include data for the total urbanised area in 1995 for Istanbul (Turkey), Lisbon (Portugal), Salvador (Brasil) and Turin (Italy). Therefore, we impute the data using two methods. Firstly, if there are other cities from the same country and year in the dataset, we take an average of the ratio between total surface area to urbanised area in the metropolitan region from the observable cities and use this to calculate the urbanised area in the missing city (the data always includes information on total surface area). If this was not possible, we used the urbanised area derived from 2002-2003 MODIS satellite data at 1 km resolution available from Schneider et al. (2003) for a given metropolitan area.

A.1.2 GDP per capita missing

The MCD 2012 dataset (UITP, 2015) is also missing GDP per capita for several cities, including; Addis Ababa (Ethiopia), Chicago, Portland OR (United States), Glasgow (Scotland), Helsinki (Finland), Izmir and Kocaeli (Turkey), Jerusalem (Israel), Johannesburg and Tshwane (South Africa), Mumbai (India), Nairobi (Kenya), Mashhad and Tehran (Iran), Melbourne (Australia), Montreal and Vancouver (Canada), Seoul (South Korea), Taipei (Taiwan), and Tokyo (Japan). We replace these missing data points using the most reliable online sources.

Furthermore, Caracas (Venezuela), Moscow (Russia), New Delhi (India) and Santiago (Chile) had no GDP data for 1995. For these cities we fill in the country level GDP per capita (in 1995

[42] An interactive chart is available at: `http://luminocity3d.org/WorldPopDen/#9/-26.2047/27.9987`.

current USD) from the World Bank national accounts data.

A.1.3 Historical GDP and population 1913 missing

Historical data in 1913 is missing for Russia, Czech Republic, South Africa, Cote d'Ivoire, Israel, Senegal and Zimbabwe. For Russia and Czech Republic, we use data in 1913 from the Former USSR and Czechoslovakia. For South Africa, GDP data is taken from the closest year to 1913 which is 1910. For all other countries, except Israel, we back extrapolate the real GDP per capita in 1913 by calculating the average growth rate over the 20 year period 1950 - 1970. Using these growth rates we calculate a rough estimate for 1913. For Israel, it is less convincing to back extrapolate as the countries growth was substantially different after 1950 as Israel did not exist before 1948. Therefore we take an average of the neighbouring states in 1913, including Egypt, Syria, Palestine and Jordan. Finally, we collect historical population data at the country level from Lahmeyer (2006).

A.2 Historical population density

We use two main datasets to calculate a proxy of population density in 1920. We collect historical data on population size and the built-up area from the HYDE3.1 dataset and the spatial extent of urban areas in 2000 from sattelite images provided by Landscan (see Goldewijk et al. (2017) and Dobson et al. (2000), respectively, for the methods used). We perform the following steps:

1. Determine the urban spatial extent of metropolitan areas in 2000, the closest year we have global satellite data from Dobson et al. (2000) using methods from the Landscan (Patterson and Kelso, 2012).
   (a) Using the Landscan dataset, we reclassify areas into urban if population density ≥ 200 pop/km^2.
   (b) We then apply focal statistics to remove highways which are classified as cells with a height and width ≤ 2.
   (c) In the resulting focal statistics raster, cells having a (height and width) value ≥ 3 are considered urban so we assign value 0 to every cell with value < 3 and 1 for very cell with value ≥ 3.
   (d) Convert raster to polygons based on cells with values 1, which results in polygons of the urban areas.

2. Overlay the spatial extent polygons in 2000 on the HYDE data from 1920 and extract the sum of the population and built-up area in 1920.

3. Divide the 1920 population by the total built-up area in 1920 within the metropolitan boundaries of a city in 2000 to obtain population density in 1920.[43]

There are three limitations with the proposed estimate. Firstly, estimates of population density in 1920 are related to the urban spatial structure in 2000, and therefore may capture some of the effect of transport technologies over the past century. For our estimates this will mean that they are downward biased as the variable may capture some of the effect of interest and it will increase the likelihood of finding that car manufacturers were present in cities with lower densities. Note however that we do not find this (see Section 4.2).

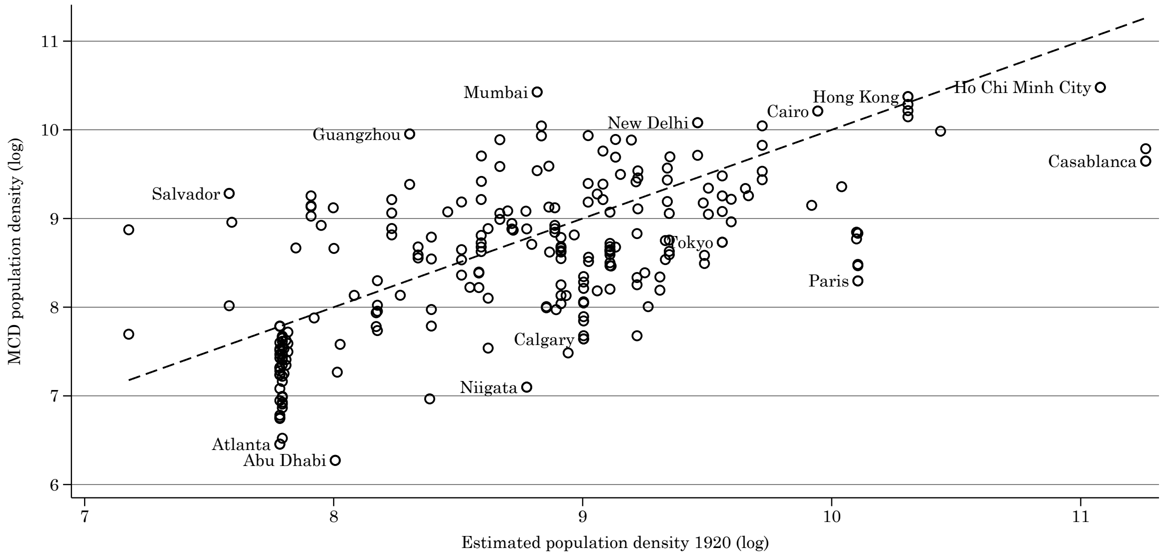
Secondly, in order to estimate population densities in the past, Goldewijk et al. (2017) require assumptions on the dynamics of population density. Using the best cross-country data available, the authors find that population density at the city level initially increases until a certain point and then decreases, similar to the findings by Kim (2007) for US cities. The authors argue that this relation can be characterised by an asymmetric bell-shaped distribution. For each country, the size and the shape of the curve differ depending on the development stage in time. While we think this is a plausible assumption, as the distribution is fitted using few data points, it may not represent the true development pattern of cities in the past and especially may be poorly suited to represent cities outside Europe and North America where most historical city level data is available.

Finally, the historical estimation procedure assumes that all cities within a country develop in the same manner according to the country level distribution, therefore the estimates are not able to capture the potential diversity in city developments over the past century and should be interpreted as an average city in a country. As our instrument is at the country level, we are interested in whether car manufacturers where more likely to be present in countries with lower city population densities, so a country level average is sufficient.

We perform various tests to confirm the reliability of the 1920 population density estimates.

[43]For three cities we were unable to compute a population density measure because the polygon did not correctly overlay the raster file of population and built-up area (Dakar and Wellington) and because the city did not exist in 1920 (Brasilia). In the case of Dakar and Wellington, we selected the grid cell adjacent to the polygon area. For Brasilia, we took the average city density of the other Brasilian cities in our dataset.

Figure A1 – Comparison of population density in data to 1920 estimate



Notes: This figure compares data from our compiled dataset which ranges between 1960 and 2012 to estimates from 1920 based on the method outlined above. The dotted line represents the 45 degree line of equality. City labels are based on minimum and maximum population densities for each 0.5 bin of estimated population density 1920 (log).

Firstly, a correlation of population density from the MCD in 1995 and our estimate using the method above for the year 2000 in logs is 0.84. This indicates a high correlation, and implies that our method of constructing population density seems valid. Secondly, Figure A1 illustrates the 1995 population density measures from MCD as compared to the estimates from 1920. The correlation in logs is 0.64 showing that density is persistent over time (Angel et al., 2010). It appears that density fell in 63% of cities between 1920 and our observatiton in the dataset which ranged between 1960 and 2012, declining on average from 8500 to 7800 people per km^2. Kim (2007) finds that average densities of US cities rose between 1890 and 1950 and fell between 1950 and 2000, however declined on average over the entire period. This is in line with Figure A1, as the majority of cities above the line of equality are in low and middle income countries where we may have expected densities to increase since the early 20th century, while cities below the line are generally in high-income countries where average densities generally fell.

Overall, it is plausible that the constructed measure captures the variation in population densities in urban areas between countries over the period, which is what we aim to measure in order to test the plausibility of our proposed instrument.

Table A1 – Commercial car manufacturers in the early 20th century

Country	Manufacturer	Founded	Closed down	Source on establishment
Australia	Holden	1914		https://archive.vn/20080322141257/http://media.gm.com/aus/holden/en/company/history/history_milestones.html#selection-847.73-877.131
Canada	McLaughlin Motor Car Company	1907	2018	http://web.archive.org/web/20080412202142/http://www.gm.ca/inm/gmcanada/english/about/OverviewHist/hist_gm_canada.html
Czech Republic	Skoda	1895		https://www.tandfonline.com/doi/abs/10.1080/00128775.1998.11648673?journalCode=meee20
Czech Republic	Praga	1907		https://www.pragaglobal.com/history/
France	Peugeot	1889		https://www.peugeot.com.au/brand-and-technology/peugeot-universe/history/cars/
France	Renault	1898		https://group.renault.com/en/our-company/heritage/the-beginning/
Germany	Audi	1910		https://www.osv.ltd.uk/brief-history-of-audi/
Germany	Benz	1885		https://www.daimler.com/company/tradition/company-history/1885-1886.html
Germany	Opel	1899		https://www.opel.com/company/history.html
Italy	Fiat	1899		https://www.lifeinitaly.com/italian-cars/fiat-history
Italy	Alfa Romeo	1910		https://www.alfaromeousa.com/a-story-that-made-history
Japan	Isuzu	1922		http://www.isuzu.co.jp/world/corporate/truck/builders01.html
Russia	Russo-Balt	1909	1918	Russian Motor Vehicles: The Czarist Period 1784 to 1917 by Maurice A. Kelly
Russia	Moskvitch	1929	2006	Cars of the Soviet Union: The Definitive History by Andy Thomson
Russia	NAMI	1927		Cars of the Soviet Union: The Definitive History by Andy Thomson
Sweden	Volvo	1927		https://www.volvocars.com/us/about/our-company/heritage
United Kingdom	Morris	1913	1983	http://www.morrisregisternsw.org/morris-the-history.html
United Kingdom	Rover	1904	2005	https://www.uniquecarsandparts.com.au/lost_marques_rover
United States	Ford	1903		https://numerov.com/dspace/es/194-id.pdf
United States	Chevrolet	1911		Chevrolet: A History from 1911 by Beverly Rae Kimes, Robert C. Ackerson

Table A2 – Additional evidence on mechanism: Private car charges in European countries, 2005

	Car manufacturer		Non car manufacturer		
	Mean	Std. dev	Mean	Std. dev	Difference (%)
Total charges petrol engine (€)	1266.00	217.63	1518.18	783.95	-19.92
Registration charges	29.00	28.61	384.68	542.02	-1226.49
VAT	326.80	28.12	345.77	117.16	-5.81
Vehicle tax	83.60	82.43	142.50	124.39	-70.45
Charges insurance	43.40	29.12	23.05	30.94	46.90
Petroleum tax	590.60	113.88	439.05	113.47	25.66
VAT on pretrolium	192.60	20.43	182.82	54.31	5.08
Total charges diesel engine (€)	1042.20	207.03	1563.00	931.79	-49.97
Registration charges	29.00	28.61	555.05	689.24	-1813.95
VAT	326.80	28.12	359.14	155.66	-9.89
Vehicle tax	122.60	127.78	238.95	221.64	-94.91
Charges insurance	53.80	36.13	28.59	38.26	46.86
Petroleum tax	369.40	105.84	252.77	50.98	31.57
VAT on pretrolium	140.40	16.68	128.55	32.97	8.44
GDP per capita indices	102.60	19.27	96.45	42.37	5.99

Notes: Data is collected from Kunert and Kuhfeld (2007). Regulatory charges are calculated for a representative European car, the Golf 1.4 with petrol engine and Golf 2.0 SDI with diesel engine in 2005. European countries with a car manufacturers in 1920 include: the Czech Republic, France, Germany, Italy, and the United Kingdom.

A.3 Car manufacturers in 1920

In Table 3 we show evidence that cities in our dataset from countries with a car manufacturer in 1920 had lower vehicle costs and more roads per capita. Kunert and Kuhfeld (2007) collect regulatory charges for a representative European car, the Golf 1.4 with petrol engine in 2005. Table A2 provides additional descriptives which indicate that, even in 2005, countries in the European Union with a historical car manufacturer, still faced far lower levels of registration taxes while having higher taxes on petroleum.

For petrol cars, total charges were 20% lower in countries with a historical car manufacturing country and registration taxes were almost zero as compared to an average of around €400 in non car manufacturing countries. Meanwhile, fuel taxes were approximately 25% higher although there is little difference in VAT on fuel. We see a similar albeit larger difference for diesel cars with total charges around 50% lower in countries with a historical car manufacturer. This provides additional evidence that car manufacturers lobbied particularly to keep costs of purchase low and encourage car ownership in their home countries, while we do not find evidence that in 2005, usage costs were particularly targeted.

Figure A2 – Histogram of key variables

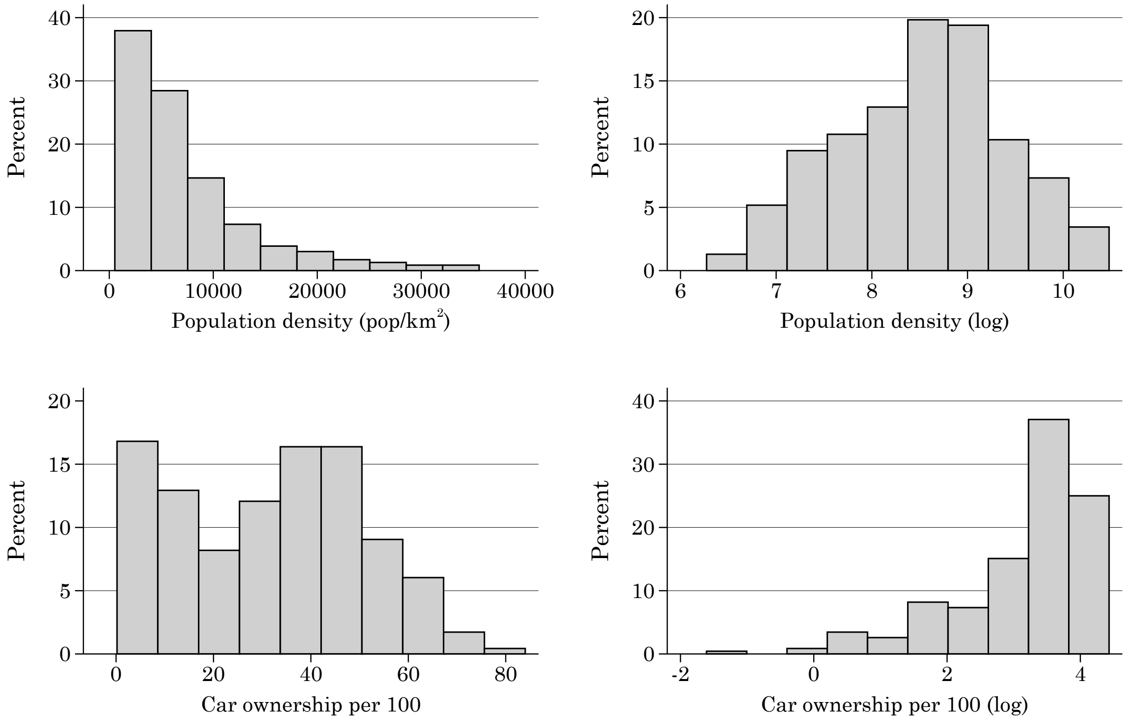


Table A3 – Additional descriptive statistics

	N	Mean	Std. dev	Min	Max
Cost of car trip	89	3.07	1.63	0.13	9.33
Annual capital car cost	92	2628.06	1517.25	152.32	10159.36
Road length per capita (m/pop)	200	3.47	3.28	0.15	17.21
Highway length per capita (m/pop)	131	0.12	0.21	0.00	1.63

A.4 Additional descriptives

In Figure A2 we report histograms of the key variables of interest. We observe that population density is approximately normally distributed when taking the log. However, it makes more sense to take car ownership in levels (as we do in the analysis); otherwise logged car ownership is strongly left-skewed.

In Table A3 we report additional descriptives for cost of car trips, annual capital costs for owning a car and length of roads and highways per capita, which we all use in Table 3. Because the cost of car trips and annual capital costs is only available in the MCD1995 data, we have fewer observations for these data.

In Table A4 we provide descriptive statistics by region and income level for population density and car ownership. Unsurprisingly, car ownership is positively correlated to the level of economic

Table A4 – Main variables by region and income level in 1995

		Population density (pop/km^2)		Cars per 100	
Region	N. cities	Mean	Std. Dev.	Mean	Std. Dev.
Africa	5	5901	3058	11.45	10.22
Asia (high income)	6	15032	10126	21.03	11.78
Asia (low/middle income)	12	18639	8774	8.11	8.18
Eastern Europe	5	7136	4204	30.60	11.38
Latin America	10	9211	3634	18.74	7.86
Middle East	7	11657	7741	13.51	7.32
North America	15	1867	752	56.79	9.36
Oceania	5	1502	529	57.54	6.14
Western Europe	35	5483	2872	41.19	10.09

Note: Calculated based on our data.

prosperity, with North America, Oceania and Europe having the highest car ownership levels. Cities in Asia have, by far, the highest population density.

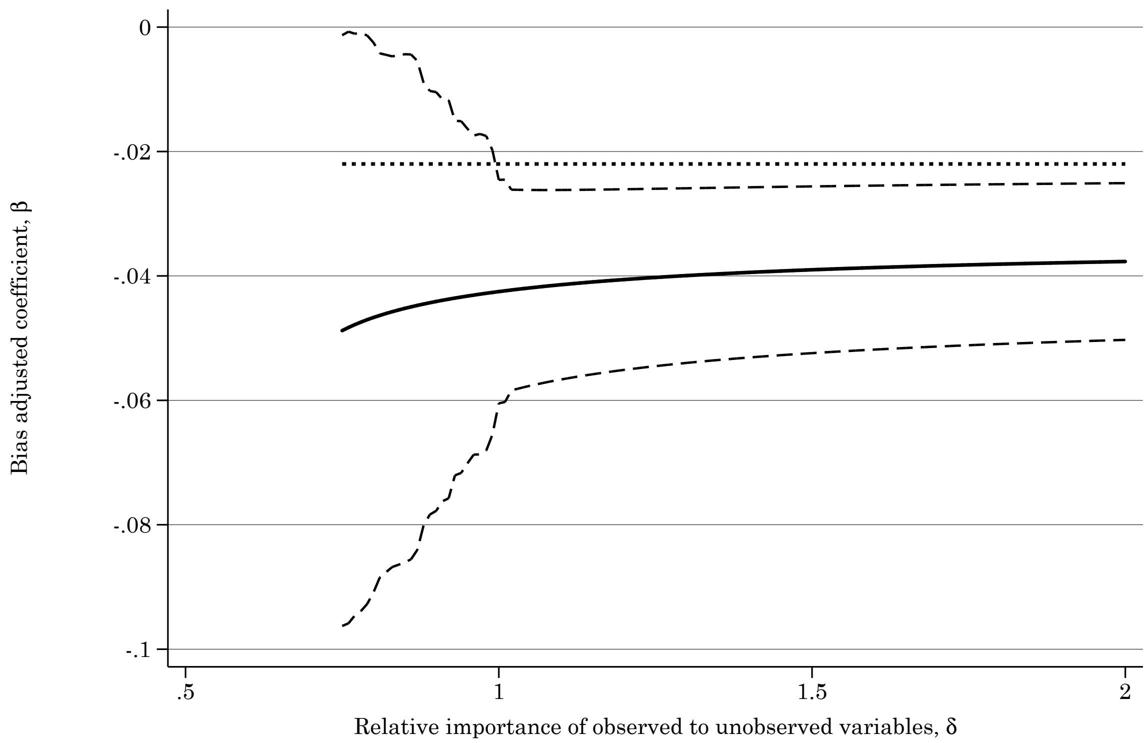
Appendix B Additional results

B.1 Oster's bias-adjusted estimates

Here we report the results of Oster's (2019) bias-adjusted estimates for different values of δ. We set R^2_{max} to 1. Recall that δ depicts the relative degree of selection on observed and unobserved variables and R^2_{max} denotes the hypothetical R^2 resulting from a regression of population density on all observable and all unobservable variables. Because of measurement error, R^2_{max} is likely lower than 1 in most empirical application (Oster, 2019). We use cluster-bootstrapped standard errors (250 replications) based on countries to construct 95% confidence bands.

Figure B1 shows that the bias-adjusted estimate is always more negative as compared to the baseline OLS estimate. This suggests that the OLS coefficient may be biased downwards and provides a conservative estimate.

Figure B1 – Oster's (2019) bias-adjusted estimator



Note: The solid line represents the bias-adjusted estimates and the dashed lines represent the 95% confidence interval where standard errors are cluster-bootstrapped (250 replications) based on countries. The short dotted line represents the OLS estimate in column (5) of Table 2.

Table B1 – First-stage results

	Dep var: Cars per 100				
	(1)	(2)	(3)	(4)	(5)
Car manufacturer 1920	27.95***	15.70***	15.94***	15.86***	14.73***
	(4.610)	(3.711)	(3.137)	(2.812)	(3.411)
GDP per capita (log)		9.318***	8.368***	8.244***	8.309***
		(0.747)	(0.729)	(1.059)	(1.081)
Pop dens. 1920 (log)			-4.406**	-4.913**	-4.524**
			(1.891)	(1.871)	(1.889)
January temperature (°C)				-0.0784	-0.117
				(0.121)	(0.132)
July temperature (°C)				0.167	0.147
				(0.175)	(0.185)
Annual precipitation (m)				-2.005	-1.875
				(2.218)	(2.095)
Altitude (km)				2.405	2.128
				(1.929)	(1.992)
Ruggedness				4.639	5.517
				(4.953)	(4.933)
French legal origin					-2.364
					(2.960)
German legal origin					-2.236
					(4.112)
Scandanavian legal origin					-2.837
					(5.104)
Decade FE	Y	Y	Y	Y	Y
R^2	0.493	0.726	0.749	0.763	0.765
First-stage F-statistic	36.75	17.89	25.83	31.80	18.64
No. of countries	57	57	57	57	57
No. of cities	123	123	123	123	123
No. of observations	232	232	232	232	232

Notes: Estimates are weighted by the number of observations per city. Robust standard errors are in parenthesis and are clustered at the country level. Statistical significance is denoted as: * p < 0.1, ** p < 0.05, *** p < 0.01. Kleibergen-Paap F-statistic is presented.

B.2 Main tables with all controls

In Table B1 and B2 we present first-stage results and IV estimates, respectively, while showing the coefficients of the control variables. As can be seen, the controls have a similar effect as in the OLS specification in Table 2.

B.3 Alternative proxies for automobile use

In Table B3 we assess the stability of our main results to various alternative measures of private vehicle usage. In column (1) we replace cars per 100 with private vehicles per 100, which therefore also includes motorbikes which are more prominent in low and middle income countries. The first-stage F-statistic is slightly lower at 14.12 (the first-stage results are reported in Table B4). One additional vehicle per 100 inhabitants results in 2.7% lower density, which is slightly larger than our main result. In column (2) we split cars and motorbikes per capita and only instrument

Table B2 – 2SLS estimates

	Dep var: Population density (log)				
	(1)	(2)	(3)	(4)	(5)
Cars per 100	-0.0293***	-0.0268**	-0.0283***	-0.0261***	-0.0247**
	(0.00718)	(0.0133)	(0.00989)	(0.0100)	(0.0102)
GDP per capita (log)		-0.0542	0.0550	-0.0103	-0.00380
		(0.146)	(0.110)	(0.109)	(0.114)
Pop dens. 1920 (log)			0.439***	0.414***	0.346***
			(0.0726)	(0.0752)	(0.0693)
January temperature (°C)				-0.00955	-0.00929
				(0.00623)	(0.00690)
July temperature (°C)				0.0119	0.00848
				(0.0118)	(0.00919)
Annual precipitation (m)				0.174	0.169
				(0.129)	(0.114)
Altitude (km)				-0.178*	-0.184**
				(0.101)	(0.0766)
Ruggedness				0.551**	0.307
				(0.240)	(0.195)
French legal origin					0.460***
					(0.150)
German legal origin					0.273**
					(0.138)
Scandanavian legal origin					-0.325
					(0.215)
Decade FE	Y	Y	Y	Y	Y
First-stage F-statistic	36.75	17.89	25.83	31.80	18.64
No. of countries	57	57	57	57	57
No. of cities	123	123	123	123	123
No. of observations	232	232	232	232	232

Notes: Estimates are weighted by the number of observations per city. Robust standard errors are in parenthesis and are clustered at the country level. Kleibergen-Paap F-statistic is presented. Statistical significance is denoted as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

cars per 100. Although it is unlikely that the coefficient on motorbikes per 100 inhabitants represents a causal effect, for example because scooters are a more attractive mode of transport in dense urban areas such as Rome and Amsterdam, the estimates in (1) and (2) along with the first-stage results in Table B4 suggest a positive correlation between motorbikes per capita and population density and a (weakly) positive correlation to car ownership. This likely explains why the coefficient of interest becomes slightly larger.

The final four columns of Table B3 regress the log of population density on the number of car km per capita and car km per car. The results in column (4), including all control variables from our preferred specification, indicates that the 2SLS estimates are statistically significant at the 1% level, although our instrument is less powerful (the first-stage F-statistic is 10). The coefficient suggests that a one standard deviation (3534 km) increase in car km per capita causes

Table B3 – Total vehicles and car km per capita

	Dep var: Population density (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Vehicles per 100	-0.0277**					
	(0.0113)					
Cars per 100		-0.0287***				
		(0.00991)				
Motorbikes per 100		0.0265***				
		(0.00736)				
Car km per capita (100 km)			-0.0204***	-0.0159***		
			(0.00338)	(0.00495)		
Car km per car (100 km)					-0.0338*	-0.0278
					(0.0198)	(0.0208)
Controls	Y	Y	N	Y	N	Y
Decade FE	Y	Y	Y	Y	Y	Y
First-stage F-statistic	14.22	21.31	8.215	12.49	1.565	1.173
No. of countries	55	55	51	51	51	51
No. of cities	123	123	107	107	107	107
No. of observations	222	222	123	123	123	123

Notes: Estimates are weighted by the number of observations per city. Controls are the log of GDP per capita, 1920 population density, January and July temperature, annual precipitation, altitude, ruggedness, and legal origin fixed effects, as in column (5) of Table 2. Robust standard errors are in parenthesis and are clustered at the country level. Kleibergen-Paap F-statistic is presented. Statistical significance is denoted as: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table B4 – Total vehicles and car km per capita (First-stage results)

	Vehicles/Cars per 100		Cars km per capita/car (100 km)			
	(1)	(2)	(3)	(4)	(5)	(6)
Car manufacturer 1920	15.02***	14.66***	44.75***	28.70***	26.99	16.18
	(3.984)	(3.176)	(15.61)	(8.122)	(21.58)	(14.94)
Motorbikes per 100		0.260				
		(0.185)				
Controls	Y	Y	N	Y	N	Y
Decade FE	Y	Y	Y	Y	Y	Y
R^2	0.707	0.779	0.365	0.726	0.0847	0.340
First-stage F-statistic	14.22	21.31	8.215	12.49	1.565	1.173
No. of countries	55	55	51	51	51	51
No. of cities	123	123	107	107	107	107
No. of observations	222	222	123	123	123	123

Notes: Estimates are weighted by the number of observations per city. Robust standard errors are in parenthesis and are clustered at the country level. Controls are the log of GDP per capita, 1920 population density, January and July temperature, annual precipitation, altitude, ruggedness, and legal origin fixed effects, as in column (5) of Table 2. Statistical significance is denoted as: * p < 0.1, ** p < 0.05, *** p < 0.01. Kleibergen-Paap F-statistic is presented.

population density to decline by around 38% in the long-run, which is essentially identical to the results from our main specification with car ownership rates as the main variable of interest. Finally, in columns (5) and (6) we estimate the effect of car km per car, however find that our instrument performs poorly and hence our 2SLS results are not very trustworthy. This may be because car manufacturers focus lobbying efforts on car ownership and purchase, while having

Table B5 – 2SLS sensitivity checks: Roads

	Dep var: Population density (log)					
	(1) Subset roads	(2) Road length	(3) Subset HW	(4) HW length	(5) Subset	(6) Roads&HWs
Cars per 100	-0.0311***	-0.0258***	-0.0277***	-0.0256***	-0.0299***	-0.0253**
	(0.0102)	(0.00953)	(0.00954)	(0.00976)	(0.00975)	(0.00998)
Road length (log)		-0.140***				-0.135**
		(0.0364)				(0.0525)
Highway length (log)				-0.0814*		0.0126
				(0.0453)		(0.0660)
Controls	Y	Y	Y	Y	Y	Y
Decade FE	Y	Y	Y	Y	Y	Y
First-stage F-statistic	16.03	15.52	18.73	17.90	18.91	17.35
No. of countries	50	50	44	44	44	44
No. of cities	108	108	94	94	94	94
No. of observations	200	200	124	124	120	120

Notes: Estimates are weighted by the number of observations per city. Controls are the log of GDP per capita, 1920 population density, January and July temperature, annual precipitation, altitude, ruggedness, and legal origin fixed effects, as in column (5) of Table 2. Robust standard errors are in parenthesis and are clustered at the country level. Kleibergen-Paap F-statistic is presented. Statistical significance is denoted as: * p < 0.1, ** p < 0.05, *** p < 0.01.

smaller effects of car usage such as fuel taxes. Still, despite weak instruments, we find a negative second-stage coefficient of car use on population density.

B.4 Roads or cars?

The earlier literature emphasises the importance of road infrastructure in the decentralisation of cities. We expect that roads and highways are complementary to car ownership, and both are required to facilitate decentralisation. Although we do not have an instrument for road length, in Table B5 we examine whether including the log of road length as a control variable meaningfully impacts our coefficient of interest.[44] If car manufacturing countries built significantly more roads and highways, it is likely that our IV results will decline as part of the effect may be captured by the roads coefficient. Column (2) indicates that including the log of road length reduces the size of the 2SLS coefficient by around 20% while the F-statistic is around the same size and the coefficient on road length is negative and statistically significant at the 1% level. Columns (4) and (6) also suggest that highways reduce the effect of cars on population density in our 2SLS estimates, however highway length appears to be less important than overall road length.

[44] Road length is measured as the total centre-line kilometres or miles of all public roads, or segregated express roads in the case of highways, therefore we may have measurement error as multiple lane roads are counted the same as single lane roads. Furthermore, several missing observations for road length and highway length is not present in the Ingram and Liu (1999) data, therefore the relevant comparison group is the subset roads and subset HW columns.

Table B6 – 2SLS Sensitivity checks: Data sources

	Dep var: Population density (log)				
	(1)	(2)	(3)	(4)	(5)
	Ing & Liu (1999)	MCD 1995	MCD 2012	Ing & Liu cities	No weights
Cars per 100	-0.0744***	-0.0347***	-0.0117	-0.133***	-0.0357***
	(0.0141)	(0.0121)	(0.0160)	(0.0500)	(0.0117)
Controls	Y	Y	Y	Y	Y
Decade FE	Y	Y	Y	Y	Y
First-stage F-statistic	25.22	14.62	6.531	5.003	20.03
No. of countries	18	51	39	18	57
No. of cities	35	100	63	31	123
No. of observations	69	100	63	50	232

Notes: Estimates are weighted by the number of observations per city. Controls are the log of GDP per capita, 1920 population density, January and July temperature, annual precipitation, altitude, ruggedness, and legal origin fixed effects, as in column (5) of Table 2. Robust standard errors are in parenthesis and are clustered at the country level. Kleibergen-Paap F-statistic is presented. Statistical significance is denoted as: * p < 0.1, ** p < 0.05, *** p < 0.01.

B.5 Sensitivity to various data sources

Our data is composed of three main data sources, obtained from Ingram and Liu (1999), Kenworthy and Laube (2001), and UITP (2015). In Table B6, we assess the sensitivity of our results to the various data sources. In columns (1) – (3), we separately estimate the effect of car ownership rates on population density, including all main controls, on each dataset. There are two important observations from this exercise. Firstly, the first-stage F-statistic appears to decline and secondly, the long-run effect of car ownership rates appears to decline. This may be due to changes over time, or due to the inclusion of additional cities and countries in the analysis. Therefore in column (4) we estimate the effect on the same cities as in Ingram and Liu (1999), but for later periods. The first-stage F-statistic goes down, while the magnitude increases, suggesting that car manufacturers were more powerful in the 1960s and 1970s in influencing policy, which is in line with anecdotal evidence. The reduction in the magnitude of the second-stage coefficient appears to be driven by the inclusion of other cities and countries. This suggests that the sample of cities in Ingram and Liu (1999) is not completely representative.

B.6 Other robustness checks

We test the sensitivity of the results to various alternative specifications and report the results in Table B7. Car manufacturers in 1920 tended to be present in higher income countries. If commercial car manufacturers were more likely to have begun in countries that had larger, more developed, markets in 1920, the instrument may be correlated to the rate of urbanisation and thereby population density in later years. Therefore, in column (1) we include the log of GDP

Table B7 – 2SLS sensitivity checks: Alternative specifications

	Dep var: Population density (log)				
	(1) Hist. controls	(2) Excl. US	(3) IV1910	(4) IV1930	(5) GDP squared
Cars per 100	-0.0334	-0.0200*	-0.0196*	-0.0333***	-0.0259**
	(0.0206)	(0.0117)	(0.0116)	(0.0110)	(0.0118)
Country GDP per capita 1913 (log)	-0.0438				
	(0.237)				
Country Pop 1913 (log)	0.0661**				
	(0.0286)				
Controls	Y	Y	Y	Y	Y
Decade FE	Y	Y	Y	Y	Y
First-stage F-statistic	10.25	12.36	10.46	25.13	10.67
No. of countries	57	56	57	57	57
No. of cities	123	111	123	123	123
No. of observations	232	205	232	232	232

Notes: Estimates are weighted by the number of observations per city. Controls are the log of GDP per capita, 1920 population density, January and July temperature, annual precipitation, altitude, ruggedness, and legal origin fixed effects, as in column (5) of Table 2. Robust standard errors are in parenthesis and are clustered at the country level. Kleibergen-Paap F-statistic is presented. Statistical significance is denoted as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

per capita and population size *at the country level* in 1913 as additional controls. The first-stage F-statistic falls to 10.25 and the estimate becomes imprecise, but the magnitude of the point estimate increases slightly and is close to our preferred specification, suggesting that our main result still holds.

Just over 10% of the observations in our dataset (12 cities and 27 observations) come from the US, which may be an outlier due to their firmly established car culture and the abundance of land for urban expansion, so in column (2) we exclude US cities. The findings suggest that the effect size decreases only slightly to −2% and becomes less precise (statistically significant at the 10% level).

In columns (3) and (4) we test the robustness to specifying the IV at different time periods (in 1910 and 1930, respectively). In 1910, only Australia did not have a car manufacturer while in 1930, both Sweden and Japan also had car manufacturers. In both cases the point estimate does not significantly change. Excluding Australia appears to reduce the strength of the instrument to F-statistic = 10 and the point estimate becomes −2% (statistically significant at the 10% level), while including Sweden and Japan increases the first-stage F-statistic to 25 and the point estimate becomes −3.3%.