Automobiles and urban density*

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

October 30, 2019

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 100 cities in 51 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 50% in the long-run. 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, R40.

---

*We thank Jeffrey Kenworthy for generously providing the data, 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 UEA Philadelphia 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 at the London School of Economics. 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.*

The invention of the automobile has profoundly influenced urban life. 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 (Glaeser and Kahn, 2004). Policy makers reacted to the widespread adoption of cars in various ways, from registration and fuel taxes to substantial infrastructure investments. As Henry Ford predicted, cars 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). Population density is the most common measure of urban form and is linked with agglomeration economies, pollution, obesity, public transport availability and sorting of the poor.[1] Therefore, studying the effect of automobiles on urban density is important as these issues are essential to the field of urban, labour, and environmental economics.

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, 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 have lower policy-induced

[1] See, for example, Ciccone and Hall (1996), Glaeser et al. (2001), Rosenthal and Strange (2004), Glaeser et al. (2008), Brownstone and Golob (2009), and Zhao and Kaestner (2010).

prices of car use and ownership through a range of mechanisms such as lower taxation and more parking. Furthermore, we will show that the presence of a historic car manufacturer is hardly (and not statistically significantly) correlated to our measure of 1920 population density, supporting our argument that historic car manufacturers are a plausible instrument for car ownership. Up to the extent that one is still concerned that omitted variables bias is an issue, we use the methodology proposed in Oster (2019) to show that our bias-corrected OLS estimates are very similar to our IV estimates.

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 higher vehicle costs tend to have higher urban densities. But as the authors acknowledge themselves, given their limited dataset and identification strategy, these results should be interpreted as suggestive.[2] Our main contribution is to improve their analysis, by adding new data and introducing a new identification approach.[3]

We apply our identification strategy to a 1995 cross-section of 100 cities from 51 countries. The results indicate that one additional car per 100 inhabitants reduces population and employment density by around 3.3% 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 60% in the long-run. Our estimates are also relevant for high-income countries with relatively low car ownership (for example Denmark) as automated vehicles will likely increase access to cars and thereby, in the absence of policy, cause cities to decentralise.

Our work is closely related to a large literature studying the effects of highways on the spatial distribution of people and jobs (Baum-Snow, 2007, 2010; Garcia-López et al., 2015; Baum-Snow

[2] Glaeser and Kahn (2004) perform their analysis using data from Ingram and Liu (1999), which contains 35 cities in 18 countries in 1960 and 1980, and instrument car ownership using legal origin.

[3] Furthermore, we control for legal origin to capture differences in urban planning regulation.

et al., 2017; Levkovich et al., 2019). This literature demonstrates that highways are an important driver of decentralisation. However, highways likely explain only a portion of car-induced decentralisation. Various other policies, such as vehicles taxes, fuel taxes, parking regimes, and socio-economic changes such as increases in income, have a strong effect on car ownership and use, and thereby on urban density, so estimates of the effect of highways arguably only give a partial view.[4] Hence, we are interested to obtain insight into the *overall effect of cars on urban density*, which is the focus of this paper.

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

2 Context and data

2.1 Literature: transportation technologies and urbanisation

From walking and public transit to cars – cities have evolved largely due to the dominant transport mode of the time.[5] Before the large-scale introduction of cars, railway development in the mid-19th century was associated with urbanisation in London and the American Midwest, but appears to have had little or no effect on population density (Atack et al., 2010; Baum-Snow et al., 2017; Heblich et al., 2018). Thereafter, the construction and subsequent expansion of the London underground system caused a large exodus of the residential population from the city of London (Heblich et al., 2018). However, subsequent development and growth of subway systems around the world only exhibited modest effects on population density during the late 20th century (Gonzalez-Navarro and Turner, 2018).

The latter finding is consistent with the new dominance of the car over the century which, more than any other transport technology, unleashed the powerful forces of decentralisation. Cars eliminated the need for fixed transportation schedules, but required substantially more space in the form of parking and roads.[6] During the mid-to-late 20th century, governments

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

[5] See Redding and Turner (2015) and Newman et al. (2016) for a discussion of this.

[6] For example, while the average space requirements per person are 2 m^2 for a moving pedestrian and 8 m^2 for a moving bicycle or bus, a car travelling about 50 kmph requires around 100 m^2 (Litman, 2019). Note that for bicycles and cars, this requirement includes parking space and assumes one occupant per travel mode.

around the world began a frenzied development of roads, both for national defence, and based on the expectation of increased car ownership and economic prosperity. This caused central-city populations to decline as roads facilitated car use and enabled people and firms to spread out. Baum-Snow (2007) documents that between 1950 and 1990, each new radial highway passing through the city centre reduces its population by around 9% and that highway construction accounts for about a third of the decline in aggregate central city populations as compared to the entire metropolitan area.[7]

Subsequent research into the effects of highways on the spatial distribution of population finds that an additional radial highway causes the central city population to decline by around 4% in Spain and China, but this effect is absent in the Netherlands, which is likely due to strict urban planing (Garcia-López et al., 2015; Baum-Snow et al., 2017; Levkovich et al., 2019). Furthermore, each new radial highway increases population growth in the periphery of a city by around $20 - 25\%$ in Spain and The Netherlands (Garcia-López et al., 2015; Levkovich et al., 2019).

2.2 Data

We use several sources of information. The most important source is the Millenium Cities Database (MCD) which contains information on a sample of 100 cities from 51 countries in 1995 and 1996 (Kenworthy and Laube, 2001; Kenworthy, 2017).[8] It contains information at the metropolitan area level on population, employment, area size, income and transportation. A key advantage of the MCD is that it uses a consistent methodology for data collection which allows us to make accurate comparisons between cities from different countries. The metropolitan area is defined as the 'commuter belt or labour market region' and hence captures the catchment area for workers in a city.[9]

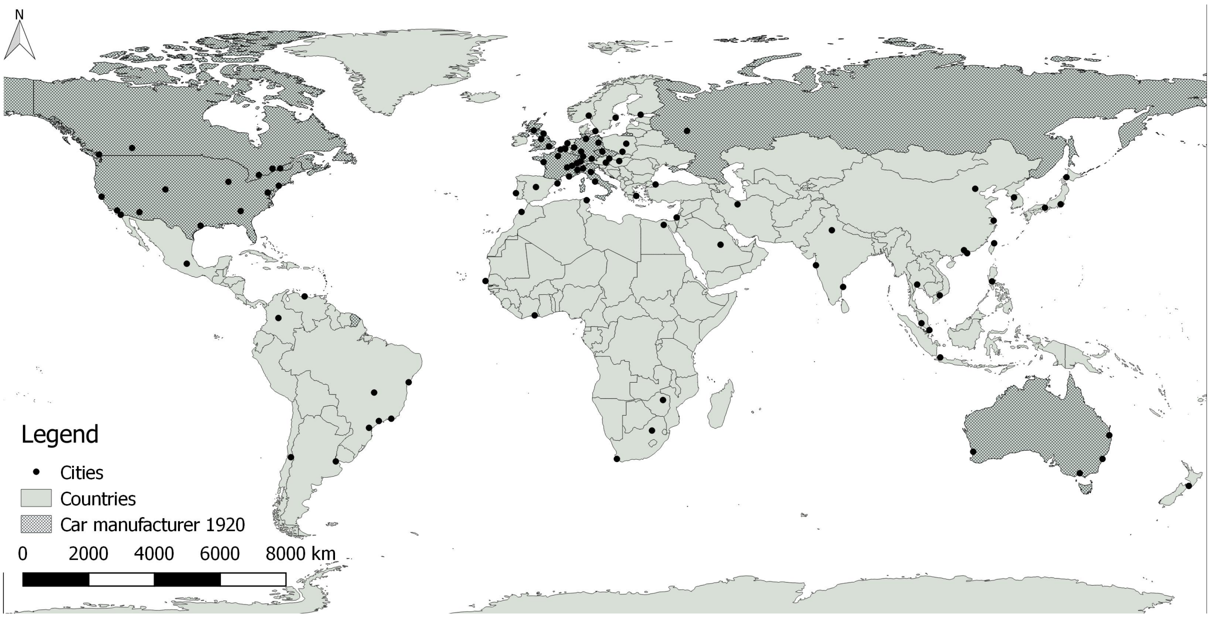
Population density is our main measure of urban structure, but we also examine other measures

[7]Comparatively, the subway effect on sprawl is about ten times smaller than the one found for highways (Gonzalez-Navarro and Turner, 2018).

[8]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.

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

FIGURE 1 – CITIES IN MCD



such as employment density and the centrality of employment. 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. The MCD also includes information on car ownership per capita, metropolitan GDP per capita and car-related variables such as the average cost of a car trip, annual capital costs of a car, highway length and the number of parking spots in the CBD.

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.1, we document which car manufacturers were present in each specific country, the year they began and (if relevant) the year they closed down, including primary sources.

To complement our historical instrument, we also collect historical data at the city level on population density and at the country level on population and GDP. Information on population density at the city level in 1920 is not available using existing sources, so we construct a measure for population density in 1920 using data from Goldewijk et al. (2017).[11] We collect historical

[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.

[11]We describe the procedure to calculate 1920 population density in Appendix A.1.2. Our data are available

Table 1 – Descriptive statistics

	Mean	Std. dev	Min	Max
Population density (pop/km^2)	7801.75	7119.15	636.14	35564.53
Employment density (jobs/km^2)	3422.64	3000.65	360.10	15127.67
Cars per capita	0.33	0.19	0.01	0.75
GDP per capita ($1000s)	19.63	14.90	0.37	54.69
Population (millions)	4.72	5.32	0.24	32.34
Total built-up area (km^2)	942.97	1392.73	64.38	10657.15
Total surface area (km^2)	4650.71	7959.18	126.09	57378.00
Built-up to surface area	0.35	0.21	0.03	0.89
Temperature summer (°C)	22.21	4.94	13.00	36.40
Temperature winter (°C)	8.58	8.98	-7.90	27.30
Car manufacturer 1920	0.41	0.49	0.00	1.00
Country Pop 1913 (millions)	55.04	96.83	0.32	437.14
Country GDP per capita 1913 ($1000s)	3.96	2.66	0.69	8.38
Pop dens. 1920 (pop/km^2)	8849.48	10449.16	1303.88	77736.80

Note: 100 observations.

information on population and GDP per capita for the year 1913, just before WWI began, from the Maddison Project Database (2018).[12]

2.3 Descriptive statistics

Descriptive statistics are provided in Table 1. Population density is around 7,800 people per km^2 and the number of cars per capita is 0.33, on average. An average city is large, containing a population of around 4.7 million and spanning a built-up area of around 940 km^2, which represents around 35% of the total surface area of the city. We put histograms of the main variables of interest in Figure A2 of Appendix A.1. They show, for example, that most cities have population densities below 10,000 people per km^2, with a median of 5,650.

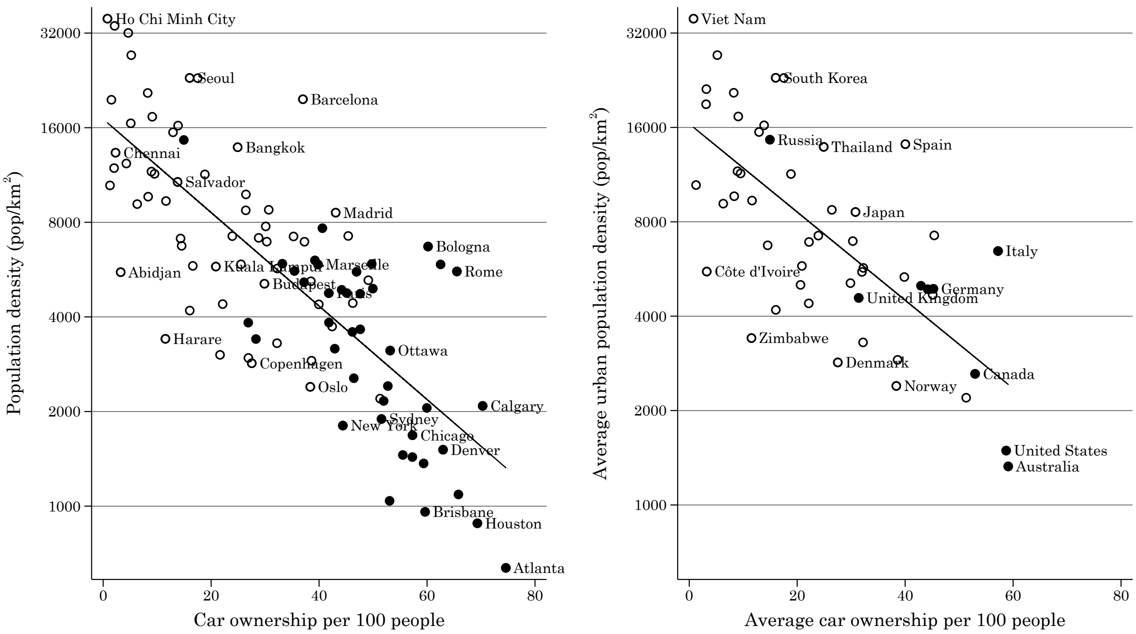
It appears that employment density is highly correlated to population density (the correlation is 0.92), so we will focus on the effect on population density. In Section 4.3, we repeat the main specifications with employment density and centrality of employment as dependent variables.[13] The left plot of Figure 2 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. It also shows that countries with historic car manufacturers have notably higher rates of car ownership and lower

upon request.

[12]Most car manufacturers were present before 1920, therefore it is more representative to use other historic information slightly prior to 1920.

[13]Centrality of employment is measured as the number of jobs in the CBD and the proportion of jobs in the CBD over the total jobs in the metropolitan area.

Figure 2 – Population density and car ownership per capita



Notes: The left plot is city level data and the right plot is at the country level. 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.

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, we set up the following regression equation:

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

where $\log(D_{ij})$ is the natural logarithm of population density, C_{ij} represents car ownership per capita, X_{ij} and G_j are vectors of observed city and country characteristics, respectively, and ϵ_{ij} is an error term. 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 $cov(C,\epsilon|X) = 0$. 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. We expect that higher incomes are correlated with higher rates of car ownership 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). Geographical factors such as the climate might also induce more or less car ownership and may be correlated to urban density, therefore we control for summer and winter temperatures. The regulatory environment and cultural factors may also play a role in determining attitudes towards car ownership and urban planning. 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 we include dummy variables for English, French, German, and Scandinavian legal origins to capture the correlation between land-use and vehicle regulations which may affect both population density and car ownership.

As it is impossible to observe all determinants of car ownership, we also perform a bias-correction approach which allows us to place a bound, denoted $\widetilde{\beta}$, on the OLS estimate of β, denoted $\widehat{\beta}$, in the presence of omitted variables. Oster (2019) shows that a consistent estimate of the bias-adjusted treatment effects can be calculated given two key parameters: (i) the relative proportion of car ownership rates explained by observables as compared to unobservables, δ, and (ii) the maximum variation in the log of population density that can be explained by observables and unobservables, R^2_{max}.[14] As the estimator can deliver multiple solutions, we need to specify how we select the most likely solution. Given that we have few controls, the bias from unobservables may bias the direction of the covariance between controls and car ownership. Therefore we select the solution closest to $\widehat{\beta}$ in case of multiple solutions.[15]

In order to tackle the issue of reverse causality, we require a long-term exogenous shock in the

[14] Oster (2019) recommends to use $R^2_{max} = 1$ as a useful upper bound, which implies that, if anything, the bias will be overstated.

[15] For $\delta > 1$ there appears always to be a unique solution.

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.[16] Legal origin may be argued to be a plausible instrument as it pre-dates the invention of the car and countries with a French civil law tend to be more regulated, hence face higher vehicle costs. 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).[17] Another issue is that the instrument appears to be weak.[18]

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.[19] 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.[20] 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 in 1995 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.[21] We present evidence and discuss the mechanism and plausibility of the instrument further in Section 4.2.

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

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

[18] Column (6) in Table A6 shows that the instrument has a first-stage F-statistic of 2.74.

[19] Our spatial unit of analysis for this variable is the country level as the presence of a car manufacturer at a less aggregate level, in particular at the city level, is less likely to be exogenous as manufacturing plants were large and employed many workers, therefore may have had a direct effect on urban structure at the local level.

[20] 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 4.4 we exclude the US as a sensitivity check.

[21] People living in cities that were more dense before the introduction of the car may have adopted fewer cars and remained more dense due to the persistence of urban structure.

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 4.4).

4.1 OLS Results

First we regress the log of population density on cars per capita. There is a strong and statistically significant negative association. Car ownership rates explain around 60% of the variation in population density between cities. One additional car per 100 inhabitants is associated with a reduction in average population density of around 3.4%.[22] In column (2) we control for GDP per capita, but the effect of car ownership hardly changes.[23]

In columns (3) to (5), we include historical population density in 1920 and some additional geographical controls. The coefficient of historic population density is positive with an elasticity of 0.39, indicating that density is persistent over time. Including 1920 population density reduces the coefficient of interest slightly to −3.0%, meanwhile there is no noticeable effect from the inclusion of weather controls in column (4).[24] 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 one additional car per 100 inhabitants is associated with a reduction in population density of around 2.6% and is still statistically significant at the 1% level.

We also present estimates from Oster's (2019) unrestricted bias correction method. Under the recommended assumption of $\delta = 1$ and $R^2_{max} = 1$, the bias corrected estimate is -0.039, with a 95% confidence interval between $[-0.052, -0.027]$.[25] We also loosen the assumptions on the δ parameter (see Figure A3 in Appendix). For $\delta \in [0.5,2]$ we calculate the bounds as

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

[23] Note that GDP per capita, and therefore income, does not have a statistically significant effect on population density *when controlling for car ownership*. 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 5).

[24] 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 5) do not indicate this to be the case.

[25] Standard errors are cluster-bootstrapped (250 replications) based on countries.

Table 2 – OLS estimates

	Dep var: Population density (log)				
	(1)	(2)	(3)	(4)	(5)
Cars per 100	-0.0342***	-0.0376***	-0.0308***	-0.0299***	-0.0268***
	(0.00524)	(0.00780)	(0.00622)	(0.00659)	(0.00604)
GDP per capita (log)		0.0691	0.0501	0.0724	0.0485
		(0.0861)	(0.0660)	(0.0591)	(0.0760)
Pop dens. 1920 (log)			0.363***	0.379***	0.286***
			(0.0692)	(0.0689)	(0.0577)
Temperature summer (°C)				0.0111	0.0109
				(0.0189)	(0.0162)
Temperature winter (°C)				0.00318	0.00454
				(0.00690)	(0.00728)
French origin					0.451**
					(0.179)
German origin					0.450***
					(0.112)
Scandinavian origin					-0.182
					(0.138)
Constant	9.744***	9.684***	6.323***	5.830***	6.330***
	(0.135)	(0.138)	(0.662)	(0.762)	(0.669)
R^2	0.605	0.609	0.693	0.699	0.761
No. of countries	51	51	51	51	51
No. of cities	100	100	100	100	100

Notes: The dependent variable is the natural log of population density. 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.

$[-0.045, -0.038]$ with an average estimated effect of -0.039.[26] This suggests that the OLS coefficient may be somewhat downward biased. We will see that bias-corrected estimates are very similar to the IV results in Section 4.3.

4.2 Car manufacturers in 1920

It is unlikely that city structure in 1920 was related to car ownership and 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. To investigate this, we will now examine how the presence of a historic car manufacturer affects transport-related outcomes at the end of the 20^{th} century.

In 1920, nine countries had a domestic commercial car manufacturer. In Table 4 we provide empirical evidence that these countries had lower ownership taxes and lower costs of car use

[26] As δ increases, the causal bound converges to around -0.038. Results from using the assumption (discussed in Section 3) that the unobservables do not bias the direction of the covariance matrix (available upon request) leads to many 'jumps' in the biased-adjusted estimates suggesting that it is not appropriate.

Table 3 – Car manufacturing countries

	N	1910	1920	1930	Examples
Australia	4	No	Yes	Yes	Holden
Canada	5	Yes	Yes	Yes	McLaughlin Motor Car Company
Czech Republic	1	Yes	Yes	Yes	Skoda, Tatra, Praga
France	5	Yes	Yes	Yes	Peugeot, Renault, Citron
Germany	6	Yes	Yes	Yes	Audi, Daimler (Mercedes), Benz
Italy	4	Yes	Yes	Yes	Fiat, Alfa Romeo
Japan	3	No	No	Yes	Isuzu
Russia	1	Yes	Yes	Yes	Russo-Balt
Sweden	1	No	No	Yes	Volvo
United Kingdom	4	Yes	Yes	Yes	Wolseley Motors, Morris, Rover
United States	10	Yes	Yes	Yes	Ford, Chevrolet, Dodge
Total	44	8	9	11	

Notes: N refers to the number of cities included in the MCD for each country. See Appendix A.1 for list of manufacturers, founding/closing dates and sources used.

Table 4 – Underlying mechanism

	(1) Car trip cost	(2) Capital cost	(3) Highway length	(4) Parking in CBD
Car manufacturer 1920	-0.337*	-0.195	0.356	0.0902
	(0.183)	(0.157)	(0.245)	(0.272)
R^2	0.629	0.688	0.652	0.453
Controls	Y	Y	Y	Y
No. of countries	43	43	39	42
No. of cities	89	92	78	84

Notes: All dependent variables are in logs. See Table A2 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, car insurance and taxes. Capital cost is defined as the annual fixed costs which includes depreciation, car insurance and taxes. Highway length is per capita. We do not observe these variables for all cities in our main sample. Controls are the log of GDP per capita and population density 1920, summer and winter temperature, and two continent fixed effects as in column (5) of Table 2. Robust standard errors are in parenthesis and are clustered at the country level. Statistical significance is denoted as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

in 1995, even when we control for GDP per capita and other controls.[27] Column (1) indicates that the price of an average car trip, which includes both variable and fixed costs, are about 30% lower in countries with historic car manufacturers. Furthermore, these countries have lower annual capital car costs (which include taxes), more highway kilometres per capita, and more parking spaces per job in the central business district, however as these coefficients are imprecisely estimated, they should be interpreted with caution.

We present the first-stage results of the instrument in Table 5. Columns (1) to (5) indicate that the presence of a commercial car manufacturer in 1920 is a strong instrument.[28] The

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

[28] Note that our variables (in particular the presence of a historic car manufacturer, GDP per capita, and historic population density) are able to explain almost all (80%) of the variation in car ownership rates between cities.

Table 5 – First-stage results

	Cars per 100				
	(1)	(2)	(3)	(4)	(5)
Car manufacturer 1920	28.17***	17.28***	17.07***	16.44***	13.62***
	(4.329)	(3.932)	(3.283)	(3.103)	(3.273)
GDP per capita (log)		8.588***	7.692***	6.959***	7.511***
		(0.975)	(0.933)	(1.245)	(1.267)
Pop dens. 1920 (log)			-5.339***	-5.798***	-5.412***
			(1.656)	(1.544)	(1.687)
Temperature summer (°C)				0.186	0.155
				(0.311)	(0.327)
Temperature winter (°C)				-0.281	-0.402**
				(0.174)	(0.170)
French origin					-1.515
					(3.437)
German origin					-4.826
					(3.036)
Scandinavian origin					-8.878*
					(4.603)
Constant	21.42***	4.668**	53.74***	58.11***	58.32***
	(2.599)	(2.126)	(14.99)	(14.51)	(14.89)
R^2	0.515	0.734	0.773	0.782	0.789
First-stage F-statistic	42.36	19.31	27.04	28.09	17.32
No. of countries	51	51	51	51	51
No. of cities	100	100	100	100	100

Notes: 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.

Table 6 – Validity test of instrument

	Historic		Current	
	(1)	(2)	(3)	(4)
Car manufacturer 1920	-0.252	0.113	-1.004***	-0.444*
	(0.291)	(0.262)	(0.275)	(0.224)
R^2	0.027	0.221	0.337	0.586
Controls	N	Y	N	Y
No. of countries	51	51	51	51
No. of cities	100	100	100	100

Notes: The dependent variable is historic (1920) or current (1995) 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, summer and winter temperature, and two continent 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.

Kleibergen-Paap first-stage F-statistic is 17.86 in the last (and preferred) specification. The instrument has the expected positive sign: countries with a commercial car manufacturer owned around 14 more cars per capita (or around 40% more than the mean city in our sample).

In Table 6 we examine whether in 1920 car manufacturers were more likely to be present in cities with lower population densities, as arguably, the instrument is more convincing if it is not (or weakly) correlated to population density in 1920. Columns (1) and (2) indicate that the

Table 7 – 2SLS estimates

	Dep var: Population density (log)				
	(1)	(2)	(3)	(4)	(5)
Cars per 100	-0.0356***	-0.0401***	-0.0394***	-0.0386***	-0.0362***
	(0.00666)	(0.0111)	(0.00922)	(0.00977)	(0.0118)
GDP per capita (log)		0.0996	0.146	0.156	0.141
		(0.126)	(0.104)	(0.0956)	(0.127)
Pop dens. 1920 (log)			0.315***	0.324***	0.239***
			(0.0672)	(0.0720)	(0.0687)
Temperature summer (°C)				0.0120	0.0112
				(0.0174)	(0.0154)
Temperature winter (°C)				-0.000487	-0.00155
				(0.00640)	(0.00882)
French origin					0.404**
					(0.194)
German origin					0.336**
					(0.162)
Scandinavian origin					-0.404
					(0.282)
Constant	9.789***	9.691***	6.790***	6.401***	6.923***
	(0.188)	(0.140)	(0.657)	(0.866)	(0.879)
First-stage F-statistic	42.36	19.31	27.04	28.09	17.32
No. of countries	51	51	51	51	51
No. of cities	100	100	100	100	100

Notes: The dependent variable is the natural log of population density. 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.

presence of a historic car manufacturer is not related to historic population density, independent of whether we include controls. To grasp the size of the estimated coefficients, we also estimate the reduced-form effect on current (1995) population density in columns (3) and (4). The results indicate that car manufacturers have a strong, negative, and statistically significant effect on 1995 population density. 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 90% confidence level.[29] Hence historic car manufacturers are not related to historic population density, but are related to current density. In the following 2SLS analysis we still control for historic population density, however the results of Table 6 indicate that the instrument is also unconditionally valid (on historic population density).

4.3 IV Results

In Table 7 we provide the 2SLS results using the presence of a historic car manufacturer as an instrument. The IV estimates are of a similar order of magnitude as the OLS estimates and a Hausman test does not reject the null hypothesis that the IV and OLS coefficients are significantly different from each other.[30] Furthermore, the IV estimate in column (5) is almost identical to Oster's (2019) bias-corrected estimate in Section 4.1. Apparently, the reverse causality issue is too small to have consequences for the estimates, at least at the city level.[31]

The preferred specification in column (5) indicates that one additional car per 100 inhabitants is associated with a reduction in population density of around 3.6%. The estimate implies that a one standard deviation increase in car ownership (19 cars per 100 inhabitants) is associated with a reduction in population density of around 50%. The effect is smaller than in Glaeser and Kahn (2004) who find an effect size around twice as large.[32]

4.4 Sensitivity

The IV results indicate that one additional car per 100 inhabitants reduces population density at the city level by 3.6% in the long-run. In this section we examine the effect on other dependent variables, perform some robustness checks, test for heterogeneous effects, and provide some tentative evidence on the middle-run effect. Results tables are presented in Appendix A.3.

4.4.1 Other dependent variables

We also present results separating population density into population and area size of the cities, and consider two additional dependent variables: employment density and a measure of employment centrality, in Table A4. Columns (1) and (2) indicate that although the effect of car ownership rates is not statistically significantly related to the log of population and size of the built-up area, the effect we find on population density appears to be via cars causing cities to spread out further.[33] This is in line with the sprawl hypothesis which argues that car

[29] The difference is -0.56, with a standard error of 0.34 and a corresponding t-statistic of -1.65.

[30] The test statistic is $\chi^2(7) = 0.71$, with a corresponding p-value of 0.87.

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

[32] Here we refer to the coefficient in Table 6, column (3), which is -0.075 or 7.2% (Glaeser and Kahn, 2004).

[33] The effect of cars on area is about the same size as we find in column (5) of Table 7, meanwhile the point estimate on population is almost zero.

ownership causes a reduction in transport costs thereby making it more attractive to build new, low density, housing at the periphery of a city which results in a drop of overall city density (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 it is not surprising that car ownership has a similar effect on firms and households.[34] Column (3) indicates that one additional car per 100 inhabitants causes a reduction in employment density of around 3.6%. There however appears to be no statistically significant effect of car ownership either on the log of the number of jobs in the CBD or the log of the proportion of jobs in the CBD (see columns (4) and (5)).

4.4.2 Other specifications

We test the sensitivity of the results to various alternative specifications and report the results in Table A5. 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 1995. Therefore, in column (1) we include the log of GDP per capita and population size *at the country level* in 1913 as additional controls. The estimate is less precise, but the magnitude is close to our preferred specification.

Many of our observations come from the US, and maybe the US is an outlier country, so in column (2) we exclude US cities. The findings suggest that the effect size decreases only slightly. In column (3) we include two continental fixed effects for Europe and the Americas/Australasia, but the results remain the same.[35] In columns (4) and (5) 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. Excluding Australia appears to reduce the strength of the instrument, however in both cases the point estimate does not significantly change.

[34] Baum-Snow (2010) also finds similar effects of highways on the decentralisation of firms and households.
[35] The base category is “Other” countries outside of these regions.

4.4.3 Heterogeneous effects

We also present results from an interaction of the fitted values from the first-stage with GDP per capita and legal origin (see Table A7 in Appendix). Column (1) indicates that the effect of car ownership on population density is not significantly non-linear with respect to income. Meanwhile, column (2) 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 may suggest that countries with French legal origins were more regulated in terms of taxation, as suggested by Glaeser and Kahn (2004), *and* planning restrictions, which may invalidate their IV approach.

4.4.4 Long-differences

We also perform a long-differences regression to provide an estimate of the middle-run effects by combining the MCD cities in 1995 with a smaller sample of cities from Ingram and Liu (1999) in the 1960s (see Table A8 in the Appendix). We observe 35 cities over an average period of 29 years and instrument the change in car ownership over the period with the presence of a car manufacturer. As we examine a shorter period of 29 years, these estimates should be interpreted as middle-run effects, as densities typically change slowly. We find evidence that the short-run effect size is about one quarter of the long-run effect, however as the sample size is small, the coefficients are imprecisely estimated. Hence, this may be a promising methodology if one is interested in the middle-run effect and better data can be obtained.

4.5 Implications

Overall, our estimates suggest that an increase in car ownership rates of one car per 100 inhabitants leads to a reduction in population density of between $3.0 - 3.6\%$.[36] We take the middle estimate of 3.3% and apply it to gauge the potential effects of growing car ownership rates in developing countries and the introduction of autonomous vehicles on urban density.

[36]See Table A5, column (2), which excludes the US and Table 7, column (5), for the smallest and largest 2SLS estimates, respectfully.

4.5.1 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 A3 in Appendix). Applying these estimates suggests that if car ownership increases to similar rates as seen in western Europe, urban density would fall by around 66% in the long-run, while if car ownership rates reach levels seen in North America and Australasia, density could fall by around 80% in the long-run.

We check these estimates for three mainland Chinese cities in our dataset: Beijing, Guangzhou, and Shanghai.[37] In 1995 average car ownership in these cities was 2.6 cars per 100 inhabitants which grew to 17.5 by 2010. According to our results this would result in a reduction in population density of around 36%, whereas the actual reduction was around 60%.[38] This suggests that changes in car ownership rates can explain up to half of the reduction in population density as our estimates represent long-run equilibrium effects and hence will overestimate changes in the short-to-middle-term.

4.5.2 Automated vehicles

Our estimates are also relevant in the broader context of future transport developments such as 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, this may cause cities to decentralise.

Fagnant and Kockelman (2015) assume that AVs are expected to increase VKT by $10 - 20\%$ in the US.[39] We consider these changes as lower an upper bounds of the effects of AVs. In scenario (A) effective car ownership increases by 10%, or 3.3 cars per 100 inhabitants, leading to a decline in population density of around 10%. In scenario (B) we consider a more extreme situation where equivalent car ownership increases by 20%, which is expected to result in population

[37] Car ownership rates and population density in 2010 is gathered from ITF (2017) and Demographia (2010), respectively.

[38] Average population density declined from 14,600 to 5,600 people per km^2.

[39] The correlation between VKT per capita and car ownership rates (in logs) at the city level is 0.94 in our dataset. We therefore equate the two measures in our counterfactual scenarios.

density declining by around 19% in the long-run.

While these estimates provide a rough indication of the potential effects of AVs, there may be reasons to think they are over- or under-estimates. On the one hand, as AVs will not need to be owned, this will likely reduce the parking space required and allow for more dense urban living, while on the other hand, because riders can engage in other activities in the vehicle, such as sleeping or working, this may 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 transport 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 causes a reduction in density of around 50% 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 urban areas increase.

This has implications for the key benefits of living and working in a city, notably agglomeration economies in production and consumption, which may justify high taxes on private vehicle ownership. 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.

Atack, J., Bateman, F., Haines, M., and Margo, R. A. (2010). Did railroads induce or follow economic growth? Urbanization and population growth in the American Midwest, 1850–1860. *Social Science History*, 34(2):171–197.

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.

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.

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

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

Ewing, R., Hamidi, S., Tian, G., Proffitt, D., Tonin, S., and Fregolent, L. (2018). Testing Newman and Kenworthys 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.

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

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.

Heblich, S., Redding, S. J., and Sturm, D. M. (2018). The making of the modern metropolis: Evidence from London. *NBER Working Paper 25047*.

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.

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.

Litman, T. (2019). Evaluating transportation land use impacts. *Victoria Transport Policy Institute*.

Maddison Project Database (2018). Maddison project database.

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.

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

Newman, P., Kosonen, L., and Kenworthy, J. (2016). Theory of urban fabrics: Planning the walking, transit/public transport and automobile/motor car cities for reduced car dependency. *Town Planning Review*, 87(4):429–458.

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.

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.

Redding, S. J. and Turner, M. A. (2015). Transportation costs and the spatial organization of economic activity. In *Handbook of Regional and Urban Economics*, volume 5, pages 1339–1398. Elsevier.

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

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.

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.

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

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.

Online Appendix

A.1 Data

A.1.1 Appending data sources

The main data source comes from the survey carried out in Kenworthy and Laube (2001) which is documented in detail in Kenworthy (2017). 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.

Istanbul (Turkey), Lisbon (Portugal), Salvador (Brasil) and Turin (Italy) lacked data for the total urbanised area and therefore urban density measures could not be calculated. We then impute the data using two methods. Firstly, if there are other cities from the same country 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.

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 current USD) from the World Bank national accounts data.

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 neighboring states in 1913, including Egypt, Syria, Palestine and Jordan. Finally, we collect historical population data at the country level from Lahmeyer (2006).

A.1.2 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.[40]

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

[40] 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.

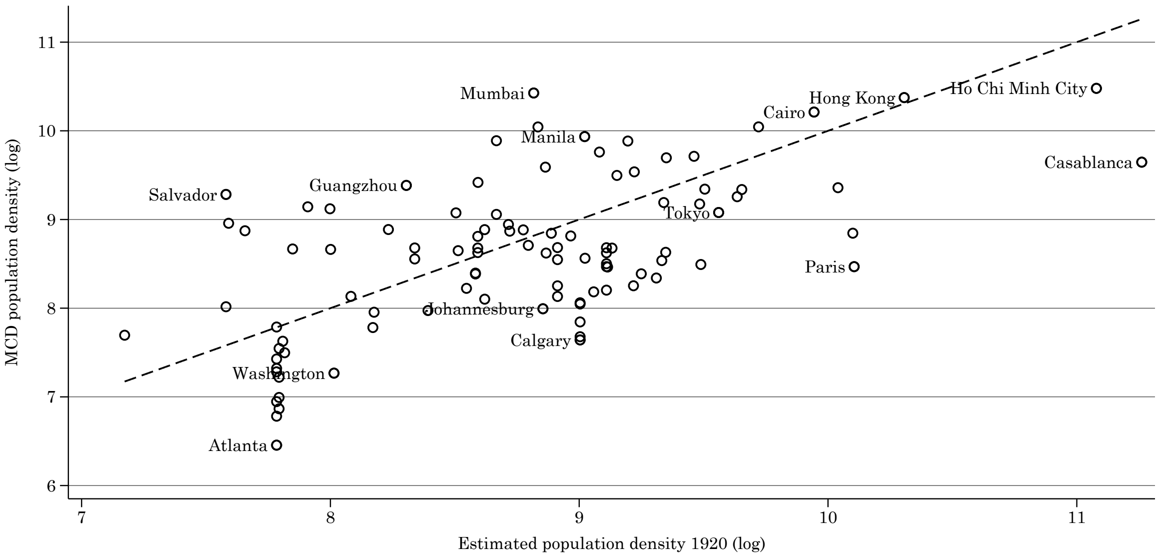
that they are downward biased as the variable may capture some of the effect of interest and it will increase the likelihood of finding that car manufacturers were present in cities with lower densities. Note however that we do not find this (see Section 4.2).

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. 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 0.61 showing that density is persistent over time (Angel et al., 2010). It appears that density fell in 60% of cities between 1920 and 1995, declining on average from 8800 to 7800 people per km^2. Kim (2007) finds that average densities of US cities rose and fell between 1890 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

Figure A1 – Comparison of population density 1995 to 1920 estimate



Notes: This figure compares data from MCD in 1995 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 MCD population densities for each bin of 0.5 estimated population density 1920 (log).

countries where we may have expected densities to increase since the early 20^{th} 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 20^{th} 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

Notes:

A.1.3 Car manufacturers in 1920

Figure A2 – Histogram of key variables

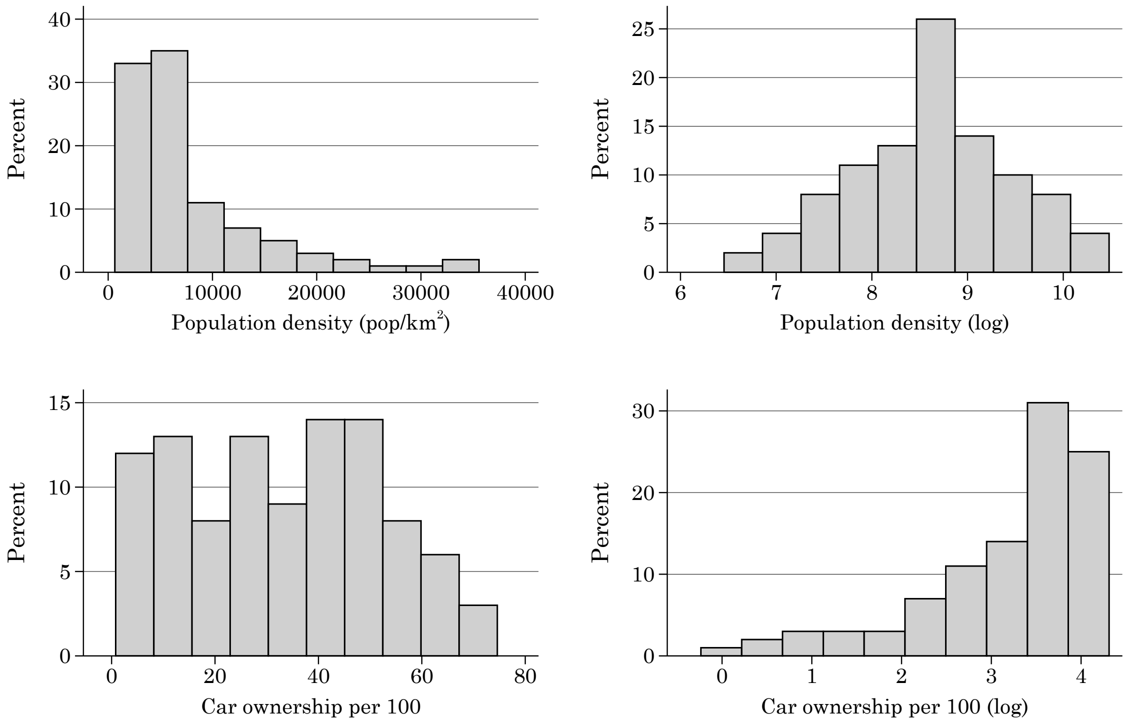


Table A2 – Additional descriptive statistics

	N	Mean	Std. dev	Min	Max
Cost of car trip	89	3.20	1.70	0.13	9.33
Annual capital car cost	92	2885.86	1707.66	152.32	10159.36
Highway length per capita	84	0.07	0.06	0.00	0.22
Parking spaces per job in CBD	84	0.29	0.30	0.00	1.88

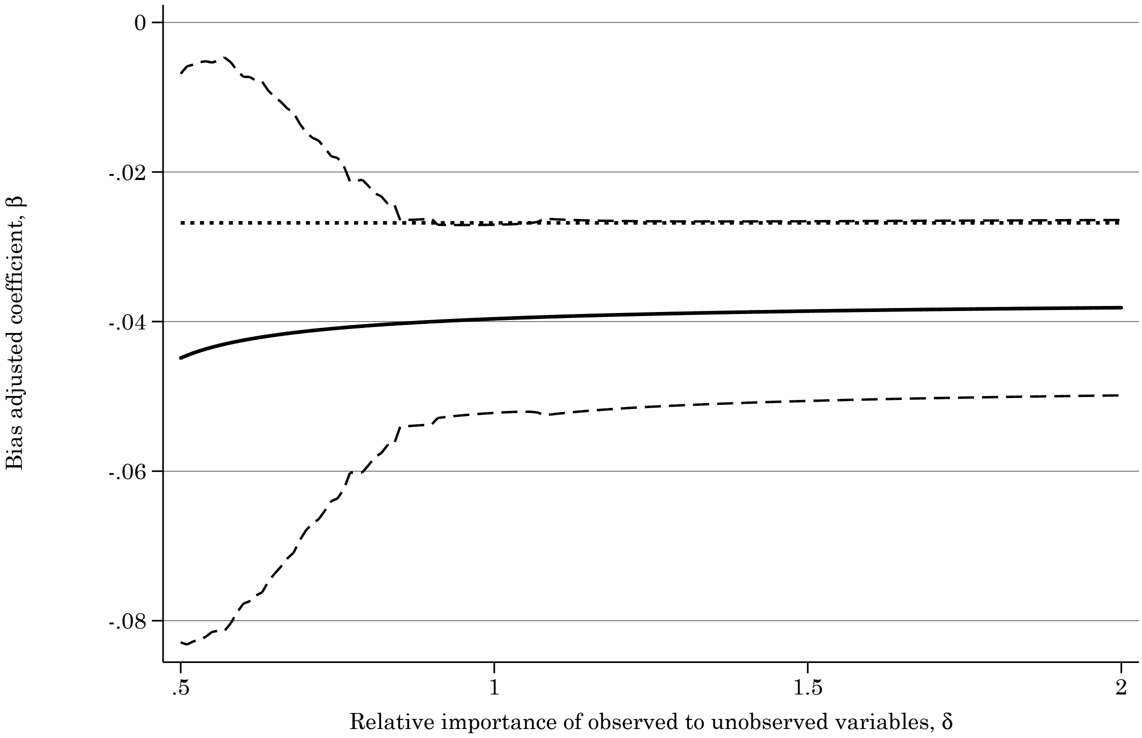
Table A3 – Main variables by continent and income level

		Population density (pop/km^2)		Cars per 100	
Region	N	Mean	Std. Dev.	Mean	Std. Dev.
Africa	5	5901	3058	11.45	10.22
Asia (developed)	6	15032	10126	21.03	11.78
Asia (developing)	12	18639	8774	8.11	8.18
Australasia	5	1502	529	57.54	6.14
Europe (east)	5	7136	4204	30.60	11.38
Europe (west)	35	5483	2872	41.19	10.09
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

Note: Calculated based on data from the MCD.

A.1.4 Additional descriptives

Figure A3 – 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.

A.2 Additional results

Table A4 – 2SLS Sensitivity checks: Other dependent variables

	(1) Population	(2) Area	(3) Emp. density	(4) Jobs in CBD	(5) Prop. Jobs CBD
Cars per 100	0.00180	0.0380	-0.0364***	-0.00600	-0.0107
	(0.0226)	(0.0267)	(0.0122)	(0.0266)	(0.0135)
Controls	Y	Y	Y	Y	Y
First-stage F-statistic	17.32	17.32	18.39	16.70	16.49
No. of countries	51	51	48	47	46
No. of cities	100	100	96	94	93

Notes: All dependent variables are in logs. 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.

Table A5 – 2SLS sensitivity checks: Robustness

	Population density (log)				
	(1) Hist. controls	(2) Excl. US	(3) Cont. FE	(4) IV1910	(5) IV1930
Cars per 100	-0.0473**	-0.0304**	-0.0362**	-0.0307*	-0.0349***
	(0.0228)	(0.0130)	(0.0164)	(0.0187)	(0.0121)
Country GDP per capita 1913 (log)	0.0290				
	(0.279)				
Country Pop 1913 (log)	0.0673*				
	(0.0350)				
Controls	Y	Y	Y	Y	Y
First-stage F-statistic	9.961	10.83	11.46	4.251	10.64
No. of countries	51	50	51	51	51
No. of cities	100	90	100	100	100

Notes: Robust standard errors are in parenthesis and are clustered at the country level. Controls are the log of GDP per capita and population density 1920, summer and winter temperature, and two continent fixed effects as in column (5) of Table 7. Statistical significance is denoted as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Kleibergen-Paap F-statistic is presented.

Table A6 – 2SLS Sensitivity checks: Robustness (First-stage results)

	Cars per 100						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Car manufacturer 1920	11.94***	12.42***	10.66***				
	(3.783)	(3.776)	(3.148)				
Car manufacturer 1910				8.556**			
				(4.150)			
Car manufacturer 1930					11.66***		
					(3.575)		
French legal origin						-11.54	-4.936
						(6.960)	(3.988)
R^2	0.816	0.762	0.814	0.756	0.779	0.082	0.732
Controls	Y	Y	Y	Y	Y	N	Y
First-stage F-statistic	9.961	10.83	11.46	4.251	10.64	2.749	1.532
No. of countries	51	50	51	51	51	51	51
No. of cities	100	90	100	100	100	100	100

Notes: Robust standard errors are in parenthesis and clustered at the country level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.3 Sensitivity checks

Table A7 – 2SLS sensitivity checks: Heterogeneity

	(1) GDP	(2) Legal origin
Cars per 100	-0.0309**	-0.0422***
	(0.0137)	(0.0120)
GDP per capita (log)	0.170	0.124
	(0.152)	(0.124)
Cars per 100 × GDP per capita (log)	-0.00380	
	(0.00411)	
French origin	0.393**	-0.294
	(0.182)	(0.258)
German origin	0.373**	0.285
	(0.167)	(0.446)
Scandinavian origin	-0.291	-0.277
	(0.282)	(0.430)
Cars per 100 × French origin		0.0220***
		(0.00667)
Cars per 100 × German origin		0.000689
		(0.0114)
Cars per 100 × Scandinavian origin		-0.00316
		(0.00874)
Constant	6.631***	6.835***
	(1.214)	(1.155)
Controls	Y	Y
First-stage F-statistic	17.32	17.32
No. of countries	51	51
No. of cities	100	100

Notes: Robust standard errors are in parenthesis and are clustered at the country level. Controls include the log of 1920 population density and summer and winter temperature. In column (1) GDP per capita (log) is re-scaled by subtracting the mean so that the coefficient of car ownership can be interpreted at the average level of GDP per capita. Statistical significance is denoted as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Kleibergen-Paap F-statistic is presented.

Table A8 – Sensitivity: Long difference

	First-difference		First-stage		IV First-difference	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Cars per 100	-0.0114***	-0.0123***			-0.00982***	-0.00669
	(0.00215)	(0.00414)			(0.00265)	(0.00705)
Δ GDP per capita ($1000s)		0.00110		0.587***		-0.00384
		(0.00430)		(0.105)		(0.00662)
Car manufacturer 1920			25.49***	13.31***		
			(3.234)	(3.203)		
R^2	0.453	0.454	0.646	0.819	0.444	0.424
First-stage F-statistic			62.15	17.28	62.15	17.28
No. of countries	18	18	18	18	18	18
No. of cities	35	35	35	35	35	35

Notes: Standard errors are in parenthesis and are *not* clustered at the country level (as there are only 18 countries). Clustering at the country level results in smaller standard errors in the second stage and larger standard errors in the first stage Statistical significance is denoted as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Kleibergen-Paap F-statistic is presented. Δ represents the difference (mostly between 1995 and 1960). The average difference is 29 years.