Supplementary material for

More connected urban roads reduce US GHG emissions

Environmental Research Letters (2017)
http://dx.doi.org/10.1088/1748-9326/aa59ba

Chris Barrington-Leigh
Institute for Health and Social Policy and McGill School of Environment
McGill University

Adam Millard-Ball
Environmental Studies Department
University of California, Santa Cruz

5 Supplementary material

Below are two supplemental sections in support of the main article. The first provides descriptive statistics for our main variables, and the second details three robustness tests for our elasticity estimates.

5.1 Descriptive statistics

Table A2 on page 34 and Table A3 on page 35 show some descriptive statistics for our key block group-level variables.

5.2 Robustness tests for elasticity estimate

In order to calculate elasticities suitable for projections, we have simplified our approach by capturing all the variation in street-network sprawl through a single measure, namely the mean degree or the fraction of 4^+ nodes. This may overestimate the road network connectivity elasticity of automobility because other aspects of urban form correlated with network connectivity, but which may vary separately and for separate reasons, are folded into our independent variable. For instance, nodal density is correlated with our connectivity measures but is in principle independent; dense urban cores and rural farm roads may both be gridded.

We have already taken two precautions to isolate the effect of interest: (1) only urban block groups are included in our regressions, and (2) the instrumental variables approach ought to isolate the effect of street connectivity. Nevertheless, as a robustness test we recalculate below the elasticities controlling for nodal density. This estimate is likely to underestimate the effect of nodal degree on automobility, because the factors which encourage higher nodal degree will also shift other metrics of urban form such as nodal density. Thus, we gain increased confidence in our estimates if they are robust to the nodal density control.

Population density is a second aspect of urban form which is likely to be correlated with our network connectivity metrics but which may be related to automobility through channels other than the connectivity itself. Regardless of streetnetwork connectivity, people living in low density areas are likely to need to travel longer distances, and are also more likely to be able to afford, and have available, a private parking space. For this reason, we also include it as a control in our revised estimates.

Table A1 on page 34 presents estimates of a model which controls for the log nodal density and the log population density in each block group. As expected, our elasticity estimates for the independent effects of the fraction of 4^+ nodes and nodal degree are attenuated slightly. Nodal density has an independent effect of the same type as our primary channel of interest, while population density is, maybe surprisingly, associated with more vehicle ownership. Notably, our primary estimate on which we rely in our later projections, changes only from -0.15 to -0.11.

	logvehicles/HH							
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)
logdegree					**−.54**†	**−.54**†	**−.55**†	**−.93**†
					(.016)	(.068)	(.016)	(.067)
logfourway	**−.090**†	**−.11**†	**−.092**†	**−.18**†				
	(.003)	(.014)	(.003)	(.014)				
logPopDensity	**.046**†	**.041**†	**.027**†	**.011***	**.049**†	**.049**†	**.031**†	**.022**†
	(.003)	(.005)	(.003)	(.003)	(.003)	(.005)	(.003)	(.003)
logNodeDensity	**−.10**†	**−.093**†	**−.074**†	**−.033**†	**−.10**†	**−.10**†	**−.071**†	**−.042**†
	(.004)	(.010)	(.003)	(.007)	(.004)	(.009)	(.003)	(.006)
Instrumented (terrain)		Y		Y		Y		Y
f.e./demean	state	state	county	county	state	state	county	county
$N_{clusters}$	2371	2371	2371	2371	2371	2371	2371	2371
Weak ID F		542		668		612		512
F	101	681	852	822	96.9	702	882	734
R^2(adj)	.238	.177	.154	.089	.241	.183	.157	.126
obs.	63360	63360	63360	63360	63480	63480	63480	63480

Significance: **0.1%**† **1%*** ***5%*** *10%*$^+$

Table A1: **Elasticity estimates with controls for nodal and population densities.** The estimates are similar to those of Table 3 on page 23, but have the added control variables.

Variable	Mean	Std.Dev.	min	max	Obs.
fourway	.20	.14	0	1	74580
logfourway	−1.86	.72	−5.2	0	74380
degree	2.7	.35	1.43	4	74580
logdegree	.99	.13	.36	1.39	74580
vehicles/HH	1.79	.38	.11	4.3	64950
logvehicles/HH	.56	.23	−2.2	1.46	64950
mean slope	5.5	5.6	0	67.6	74560
fraction >10°	.15	.20	0	.99	74560

Table A2: **Means and standard deviations of key variables**

One other concern with our main estimate is spatial dependence. If street connectivity, car ownership and travel in one block group are affected by those same variables in nearby block groups or by other unmeasured variables which are correlated across nearby block groups, spatial autocorrelation may lead to bias in our estimates. Even though our standard error estimates are clustered at the county level, this may not capture more local-level spatial dependency, nor dependency across county lines. We therefore consider a more general model which allows both for spatial lags in the dependent variable and spatial autocorrelation in the error term. Conceptually, the local measure of automobility A for a given block group is now

$$A = \rho\, w_1 \cdot \tilde{A} + \beta_0 X_0 + \beta_{\mathbf{1}} \cdot \mathbf{X}_1 + \mu \quad (7)$$

$$\mu = \lambda\, w_2 \cdot \tilde{\mu} + \xi \quad (8)$$

where $\tilde{A}$ is a vector of the automobility measure in nearby block groups, w_1 is a

	logfourway	degree	logdegree	vehicles/HH	logvehicles/HH	mean slope	fraction >10°
fourway	**.90**†	**.85**†	**.83**†	**−.35**†	**−.35**†	**−.22**†	**−.23**†
	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$
logfourway		**.81**†	**.80**†	**−.35**†	**−.35**†	**−.24**†	**−.25**†
		$<10^{-5}$	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$
degree			**1.00**†	**−.38**†	**−.38**†	**−.29**†	**−.30**†
			$<10^{-5}$	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$
logdegree				**−.38**†	**−.37**†	**−.29**†	**−.31**†
				$<10^{-5}$	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$
vehicles/HH					**.98**†	**.11**†	**.11**†
					$<10^{-5}$	$<10^{-5}$	$<10^{-5}$
logvehicles/HH						**.11**†	**.11**†
						$<10^{-5}$	$<10^{-5}$
mean slope							**.94**†
							$<10^{-5}$

Table A3: **Raw correlations between key variables measured at the blockgroup level**

spatial weighting vector, ρ is the coefficient for the spatial lag effect, and the X's and β's are as before. The unmodeled component, or error term, of observed A for the given block group is μ, and it is related to the error $\tilde{\mu}$ in nearby block groups through the coupling vector w_2 and the spatial autocorrelation coefficient λ. The remaining idiosyncratic error is the spatially independent error term ξ. This model is known as a combined spatial-lag (for ρ) and spatial error (for λ) model. We estimate this specification for our data using, as before, an instrumental variable in which X_0, our measure of road connectivity, is first projected on the instruments and X_1 before being used to estimate A.

We estimate this IV spatial-lag spatial-error model using the estimator implemented in the PySAL software package's *GM-Combo* routine [6]. In general, spatial weights matrices are imposed based on reasonable assumptions, and we use the same symmetric matrix $\mathbf{W}$ to account for all the linkages w_1 and w_2.

We experiment with three commonly used spatial weights matrices: 10-nearest neighbors, and inverse distance weights within a 2km and 5km band respectively. In all cases, our estimated elasticities are similar to those in Table 3 on page 23, and lie between –0.18 and –0.15. The standard errors are smaller than those in Table 3 on page 23. We conclude that while theoretically an issue, spatial dependence does not influence our estimates in practice.

A third line of robustness testing[12] relates to the linearity of the relationship between $\log\left(D^{4+}\right)$ and the log of vehicles/household. In order to test for variation in this elasticity across different types of neighborhoods, we separate our sample into three terciles according to the value of D^{4+}. Within each sub-

[12] We thank a reviewer for suggesting an investigation along these lines.

	Fourway (D^{4+}) range		
	0.00–0.12	0.12–0.21	0.21–1.00
	(1)	(2)	(3)
logfourway	**−.21**†	*−.47*$^{+}$	**−.18***
	(.060)	(.27)	(.055)
IV	Y	Y	Y
N_{clusters}	1562	1934	2151
Weak ID F	90.8	47.1	377
F	13.8	20.8	21.5
R^2(adj)	−.037	.033	.139
obs.	19870	21170	22320
log likelihood	5008	3972	469

Significance: **0.1%**† **1%*** ***5%*** *10%*$^{+}$

Table A4: **Elasticity estimates for different street-network sprawl regimes.** The estimates are similar to those in Table 3 on page 23 but are calculated for subsets of the sample, separated into terciles of street connectivity (D^{4+}).

sample, we estimate the IV model, as in Table 3 on page 23. The elasticities are not individually statistically different from that calculated for the whole sample, and the large standard errors, particularly with the estimate for the central tercile, mean that there is no statistically significant trend across the three terciles. Also, the point estimates for elasticity in each of these smaller samples is at least as high as our primary estimate used below for projections. Table A4 on page 36 presents these results.