Appendix - For Online Publication

Infrastructure Inequality: Who Pays the Cost of Road Roughness?

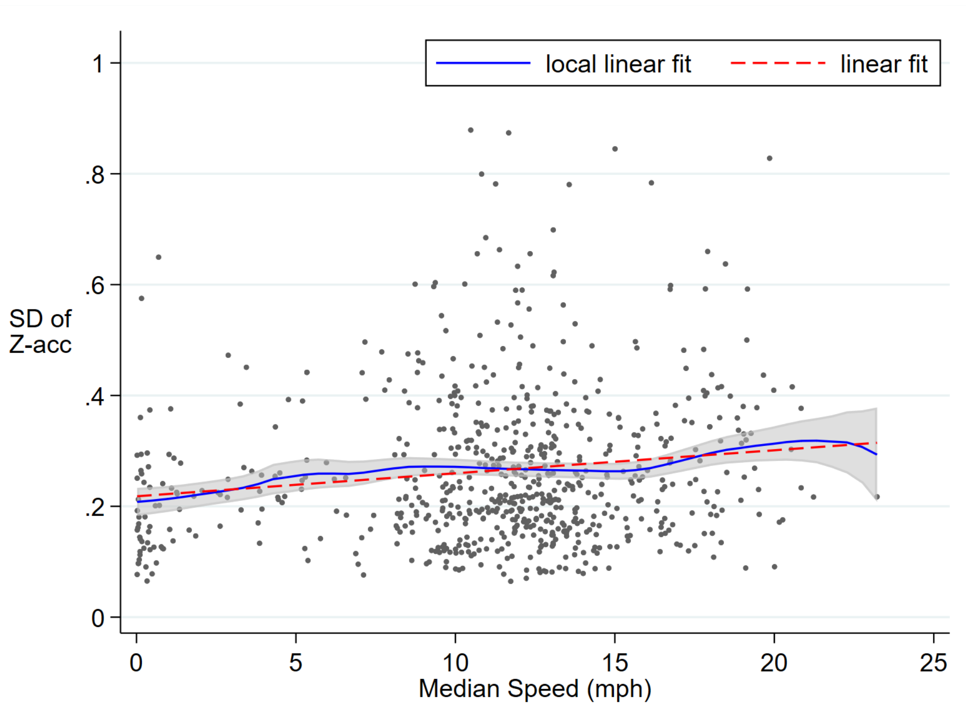
Lindsey Currier, Edward L. Glaeser and Gabriel E. Kreindler

December 15, 2023

Contents

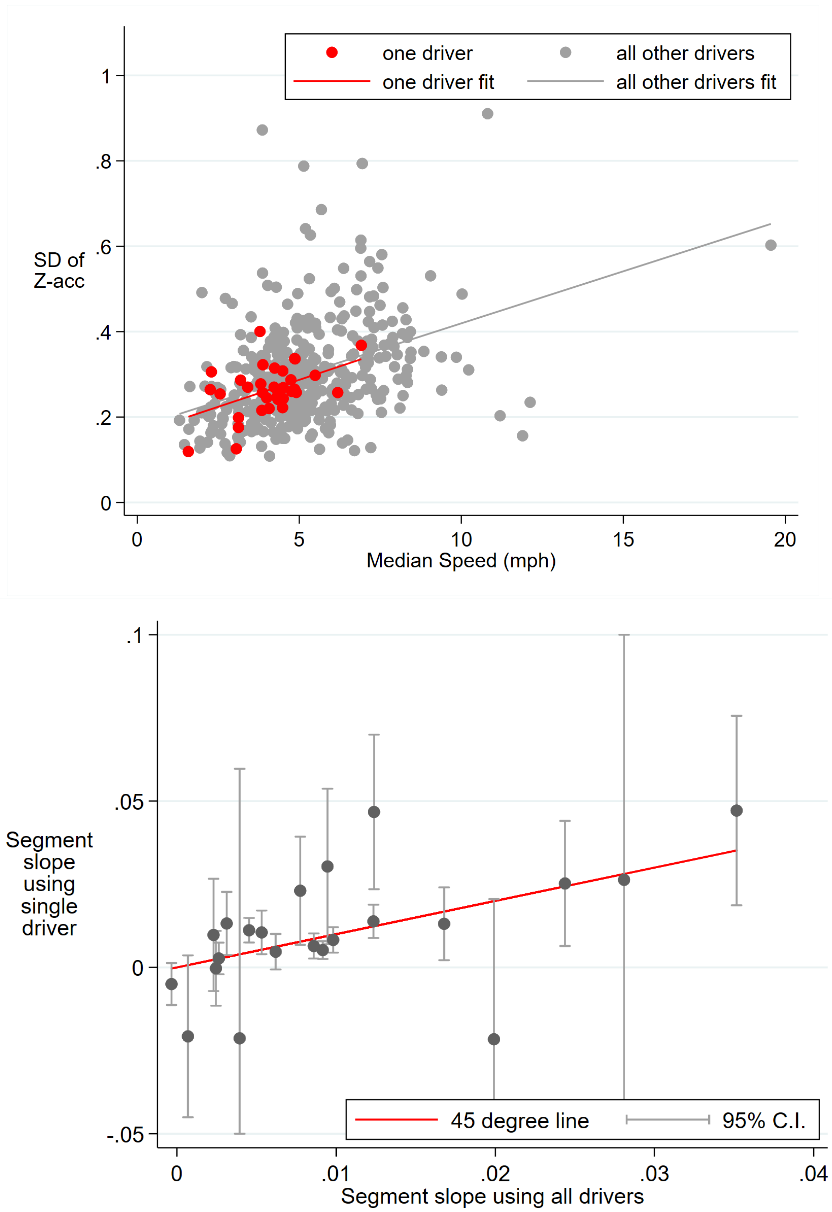
A Appendix Figures

Figure A.1: Linear and Local Linear Speed Slopes



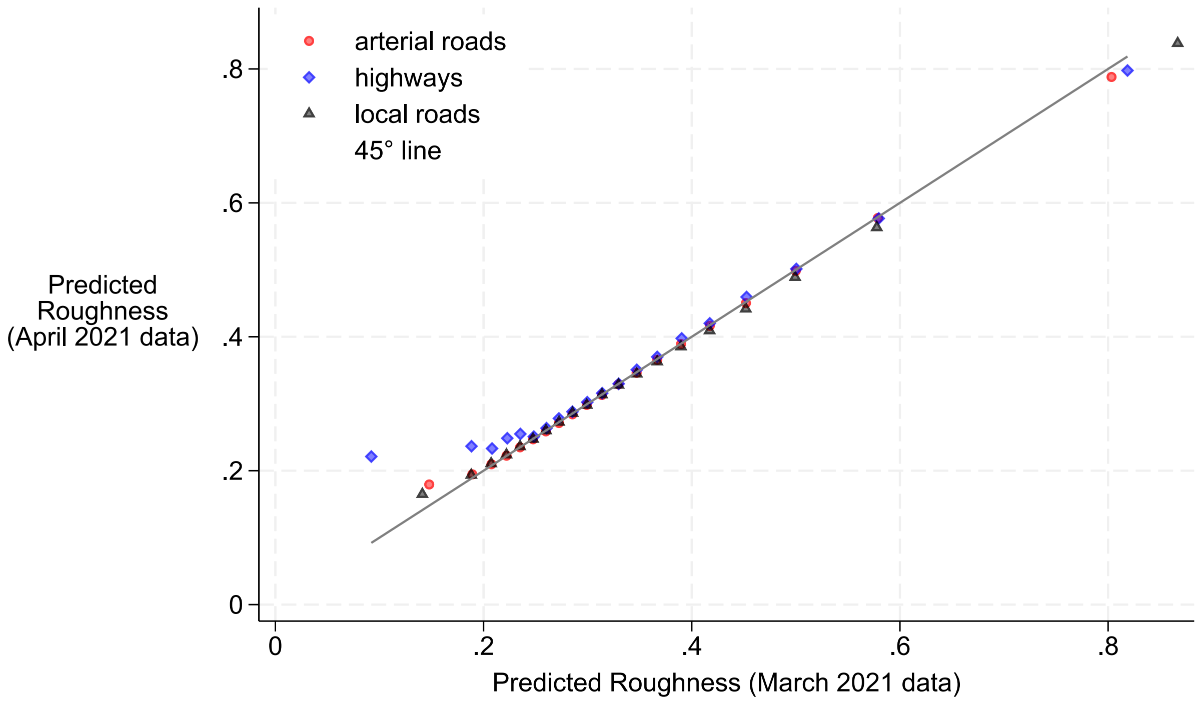
Note: This graph shows the relationship between the standard deviation of Z-acceleration and median speed. Each observation is at the driver-segment level. The sample is the road segment with the largest number of observations in April 2018 (after we keep no more than 33 observations per road segment per day). The blue, solid line is a local linear fit with 95% confidence intervals in gray. The red, dashed line is the linear fit. This graph suggests that the relationship between speed and the standard deviation of vertical acceleration is approximately linear.

Figure A.2: Segment- and Driver-specific Speed Slopes



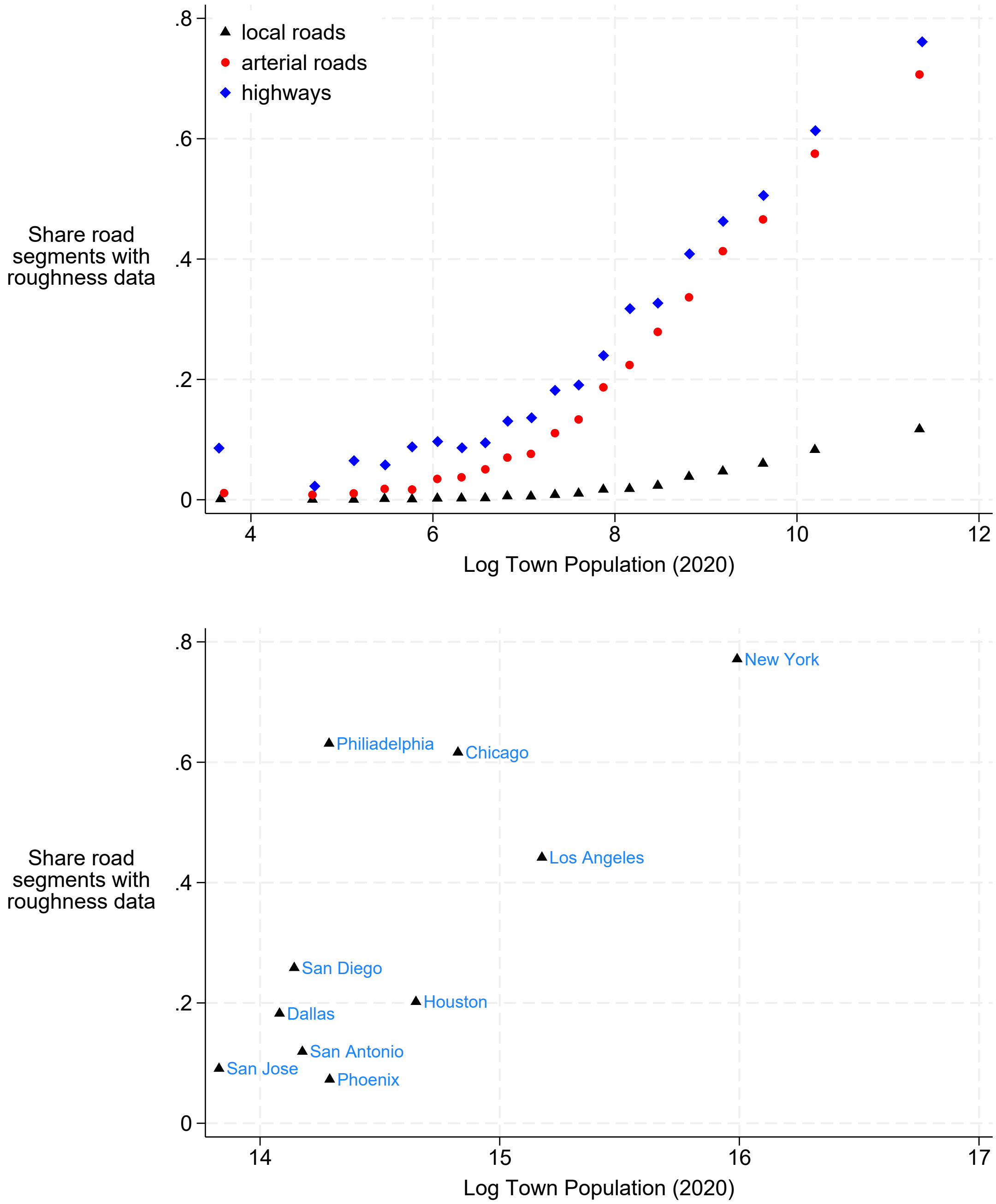
Note: These graphs show that the relationship between median speed and standard deviation of Z-acceleration is for the most part not driver-specific within segment. In the top graph, we select the driver with the most observations over any given segment. The red points represent the 34 observations for the selected driver and segment (linear fit in red) and the gray points are all the other observations for that segment (linear fit in gray). (Results are virtually identical when dropping the outlier.) In the bottom graph, we select all segments with at least 20 observations from the same driver. For each such segment, we compute the segment-specific slope excluding that driver (X axis), and the segment-driver-specific slope using only the observations for that driver (Y axis). The gray lines represent 95% confidence intervals of the segment-driver-specific slope (censored at -0.05 and $+0.1$). The 45 degree line in red falls within nearly all the confidence intervals of the segment-driver-specific slopes.

Figure A.3: Cross-Validation of Segment-Level Uber Road Roughness (Chicago)



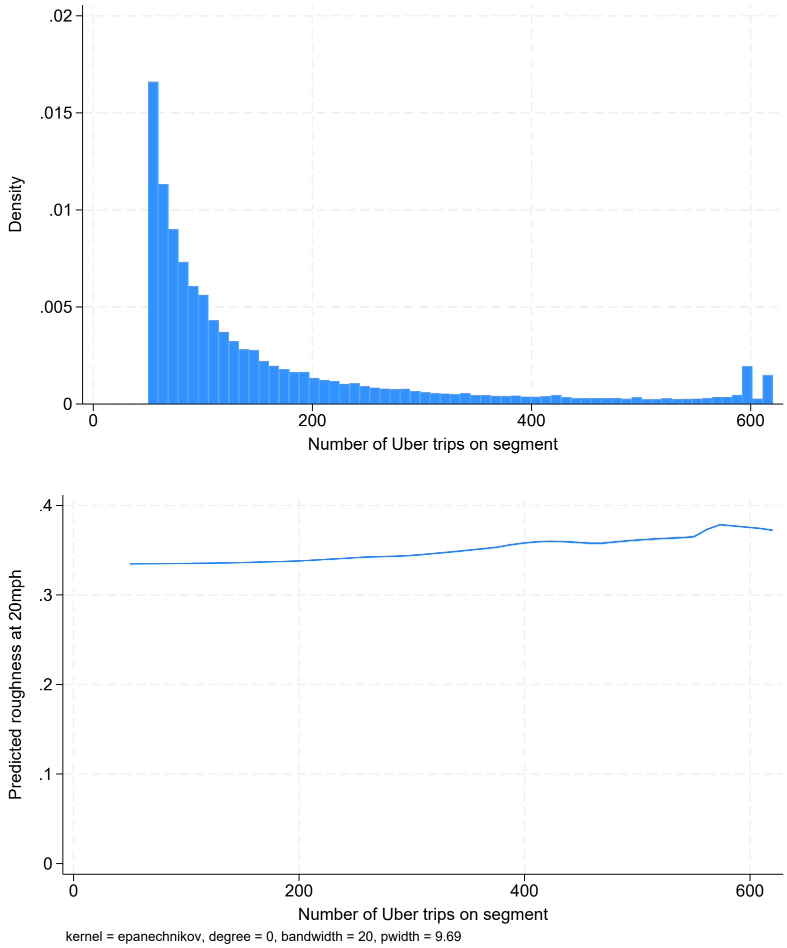
Note: This graph shows the binned scatter plot of predicted roughness at the road segment level in Chicago, using only March 2021 data (X axis), and using only April 2021 data (Y axis). Predicted roughness is computed using 20mph for local roads, 32mph for arterial roads, and 48mph for highways. See Table A.3 for additional results..

Figure A.4: Uber Measure Coverage and Population (Town level)



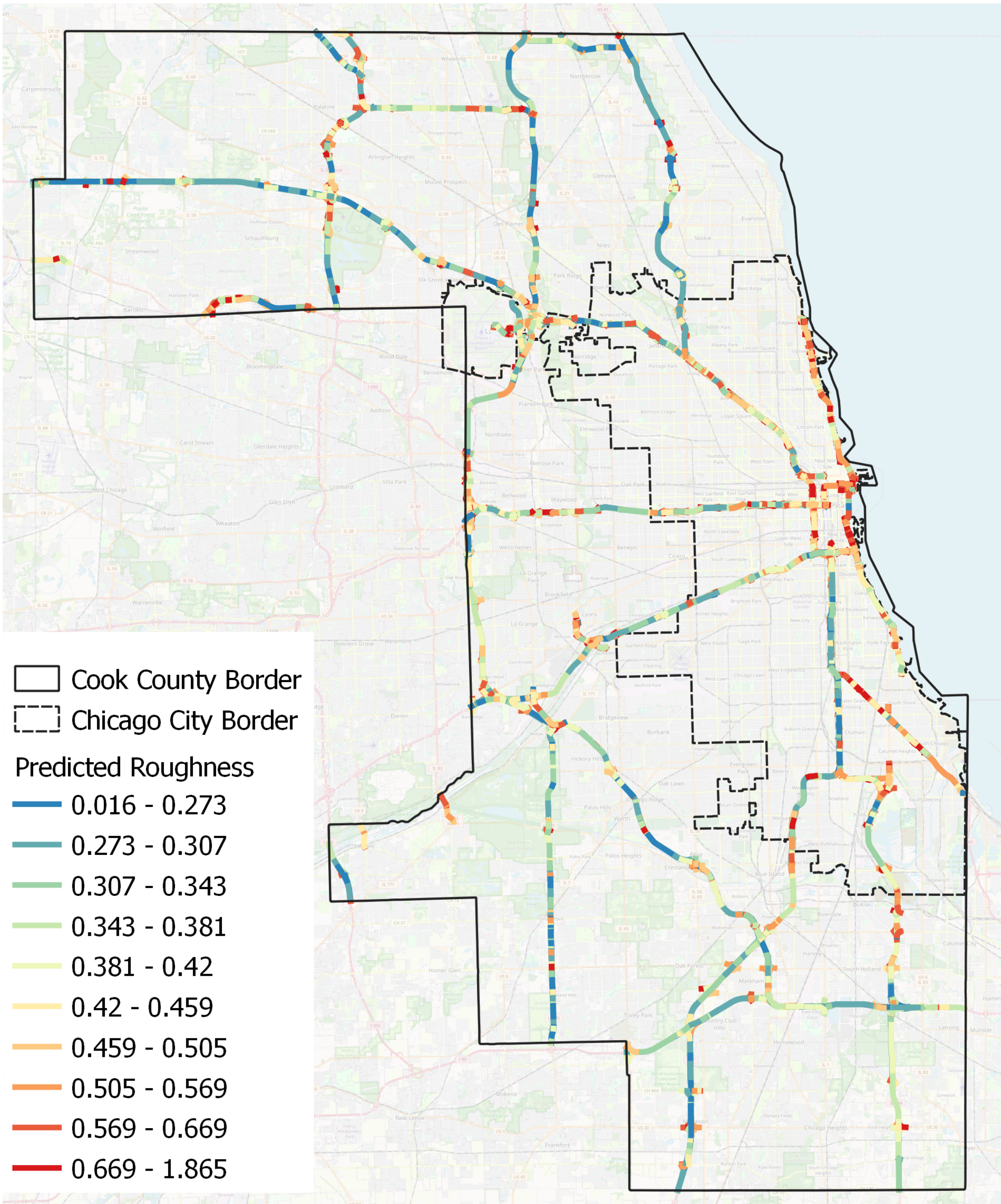
Note: The top graph shows a binned scatter plot of the share of road segments with Uber road roughness data at the town level, as a function of log town population, by road type. The bottom graph displays coverage for local roads for the census places with census population above 1 million. See Table A.1 for additional results.

Figure A.5: Number of Uber Trips per Segment and Road Roughness (Local Roads)



Note: The top graph shows the histogram of number of Uber trips per segment (censored at 620, which corresponds to 20 trip per day in our one month of data) in our sample of 1,829,526 local road segments. The bottom graph plots a locally linear regression of predicted road roughness at 20 mph versus trips per segment..

Figure A.6: Predicted Roughness at 48 mph in Cook County, IL

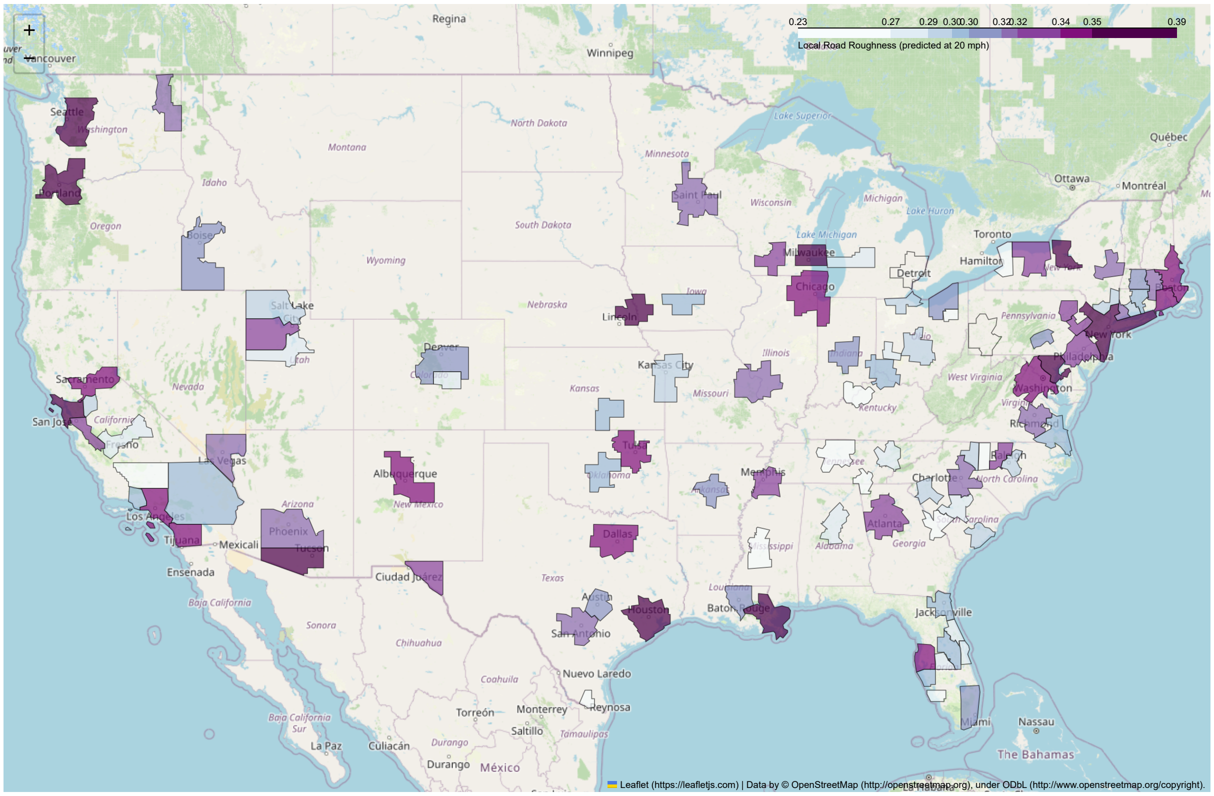


Note: This map plots predicted road roughness for all highway road segments in Cook County. Colors correspond to deciles of the roughness distribution at 48mph.

Figure A.7: Brick Roadways in the Village of Wilmette, IL

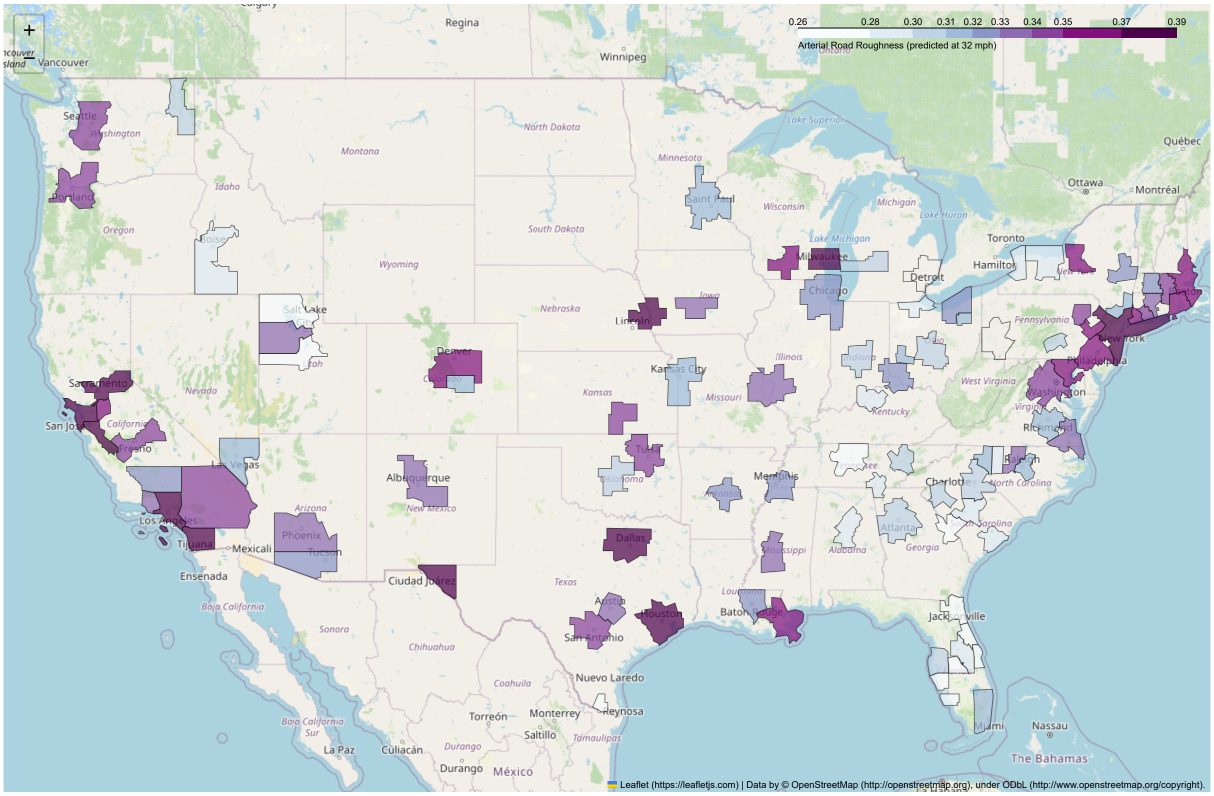
Note: The top panel highlights the high predicted roughness in the village of Wilmette in New Trier Township in Cook County, IL. (Colors correspond to deciles of the roughness distribution.) The bottom panel shows brick roadway in a Google Street View image (©Google).

Figure A.8: Local Road Roughness at MSA level (top 100 MSAs)



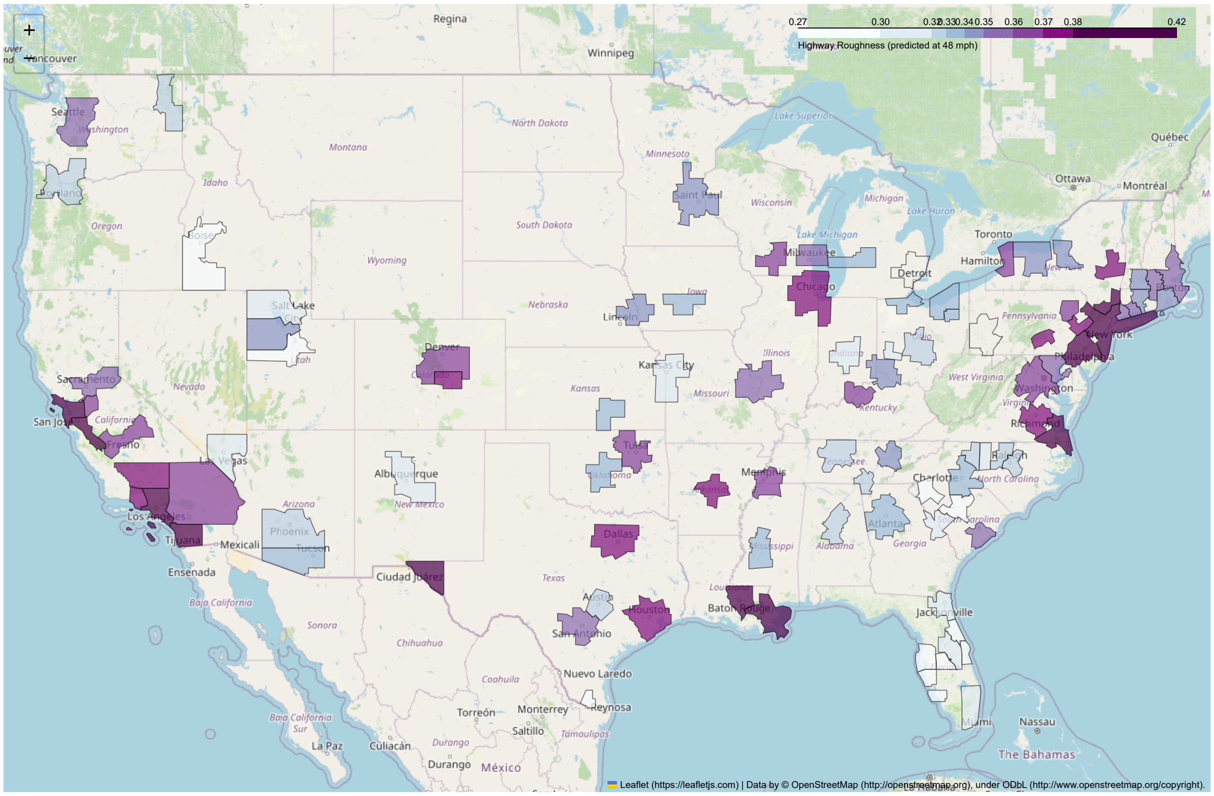
Note: This map plots the average road roughness on local roads for the top 100 MSA by population in our data. MSA-level average road roughness is winsorized at 2.5%..

Figure A.9: Arterial Road Roughness at MSA level (top 100 MSAs)



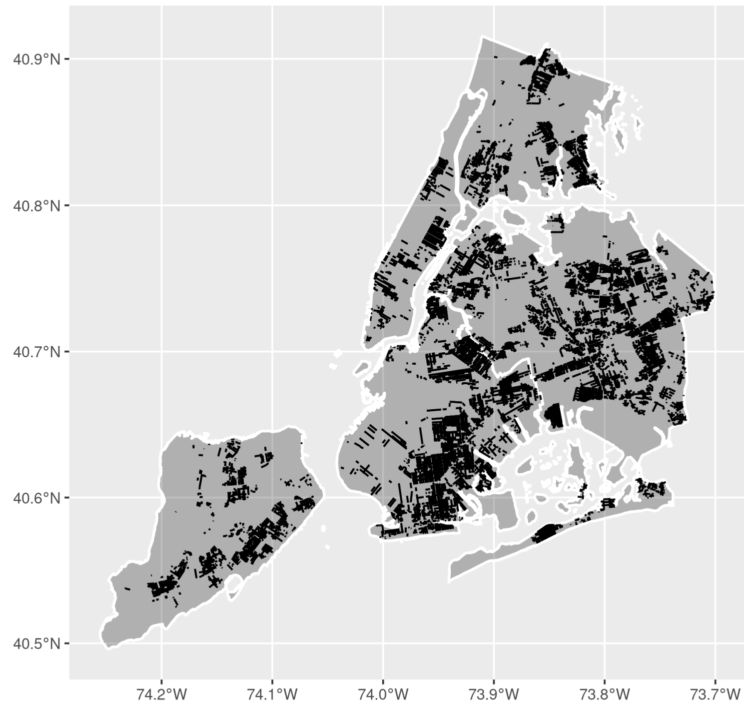
Note: This map plots the average road roughness on arterial roads for the top 100 MSA by population in our data. MSA-level average road roughness is winsorized at 2.5%..

Figure A.10: Highway Roughness at MSA level (top 100 MSAs)

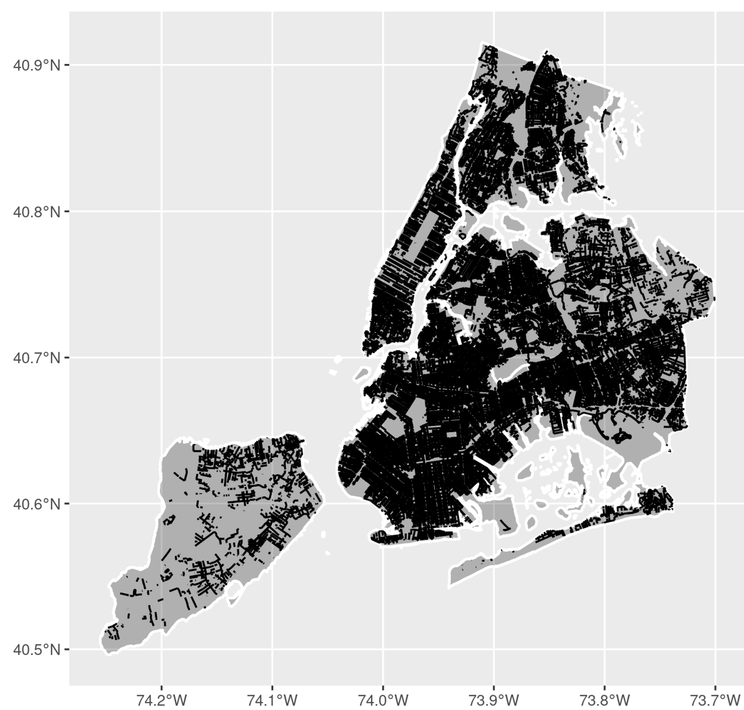


Note: This map plots the average road roughness on highways for the top 100 MSA by population in our data. MSA-level average road roughness is winsorized at 2.5%..

Figure A.11: Uber and PCI data coverage on local roads in NYC



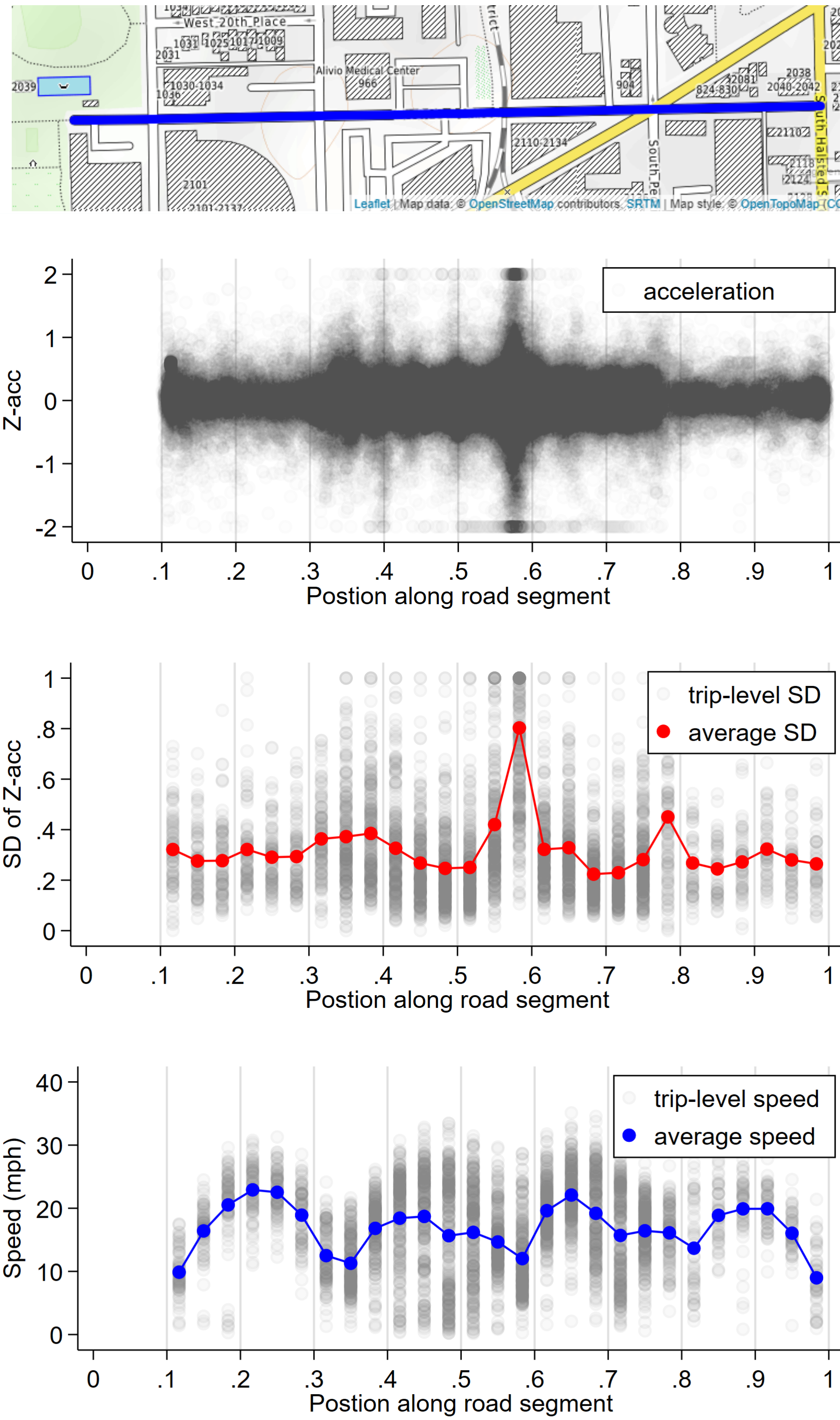
(a) PCI (June-October 2021)



(b) Uber (August 2021)

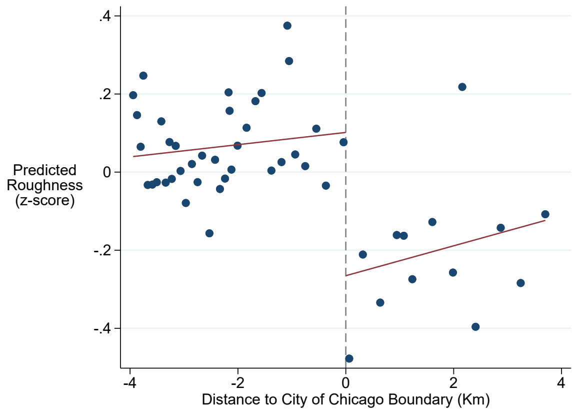
Note: These figures show the local road segments with data for PCI and Uber roughness, respectively. PCI from New York City from June-October, 2021. Uber data is for August alone..

Figure A.12: Uber Data Has Signal: Railroad Crossing

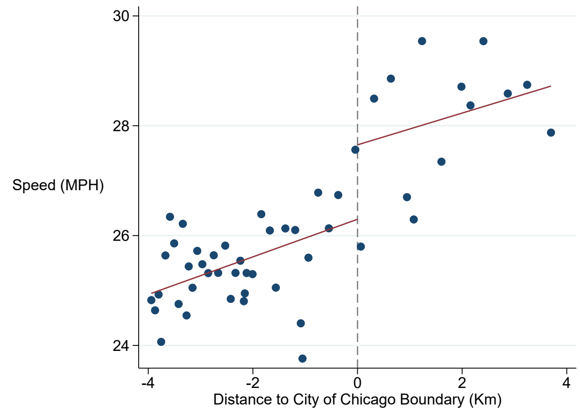


Note: These graphs show vertical acceleration, its standard deviation, and speed, for all trips covering a given road segment with a railroad crossing in Chicago (shown in blue in the map in the top panel)..

Figure A.13: Predicted Roughness and Speed around the Chicago border



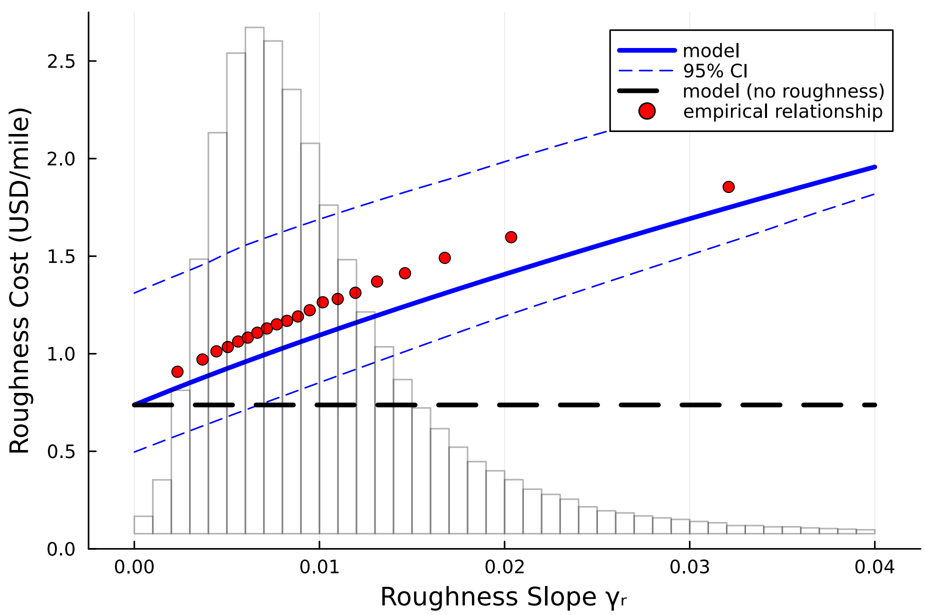
Panel (A) Predicted Roughness



Panel (B) Speed

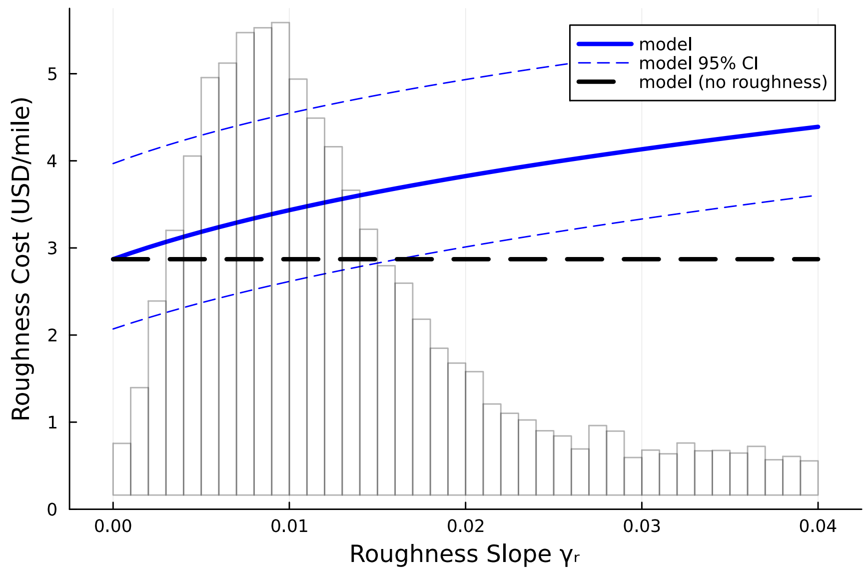
Note: Version of Figure 8 for artery road segments..

Figure A.14: Total User Cost of Roughness vs No Roughness (Local Roads)



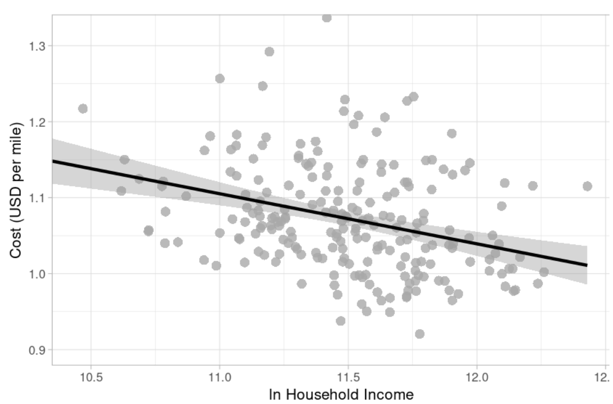
Note: This graph replicates Figure 9, adding the binscatter plot of total user costs computed using actual speeds measured in the data (red dots)..

Figure A.15: Total User Cost of Roughness vs No Roughness (Local Roads, Chicago Estimates)

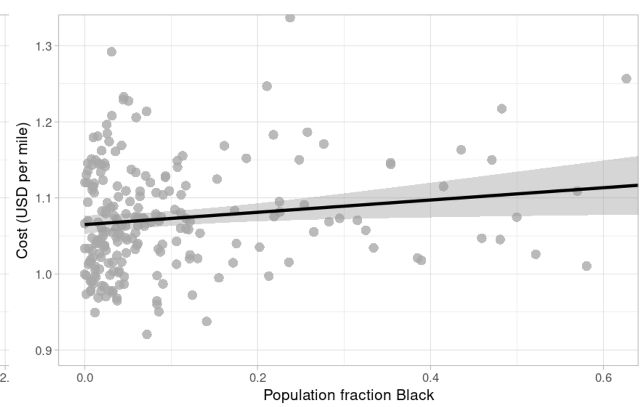


Note: This graph replicates Figure 9 for the Chicago road resurfacing sample and using the estimates from Table 7. At the median level of road roughness in this sample, the cost of road roughness (the blue line minus the dashed black line) is 0.62 USD. Also at the median level of road roughness, the cost of an additional one standard deviation of road roughness is equal to 0.44 USD, with a bootstrapped 95% confidence interval of $[0.43, 0.45]$ USD.

Figure A.16: Road Quality in the New York-Jersey City Metro Area



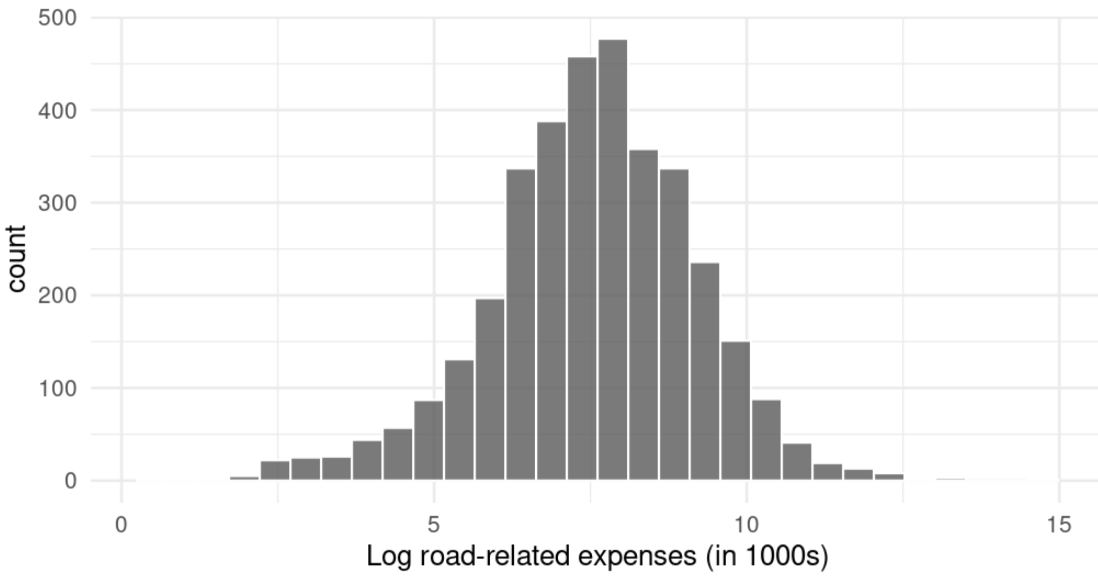
(a)



(b)

Note: This figure shows the average per mile cost of local roads by log household income and percentage of the population that is Black in panels (a) and (b), respectively, for towns in the New York-Jersey City MSA..

Figure A.17: Road Expenditure by Local Goverments



(a)

Note: This figure shows a histogram of 2017 spending on local roads for the towns/cities/CDPs matched to the ASLGF with positive spending. Spending at the country level is assigned to places based on the share of overlapping land area. We show results under a log(x) transformation because our data contains a long right tail. For example, NYC reported 2.5 billion, Chicago reported 670 million, Los Angeles 550 million, and Seattle 430 million. The median town in the analysis reported spending 2 million.

Figure A.18: Dallas grids with greater than zero Uber coverage on local roads

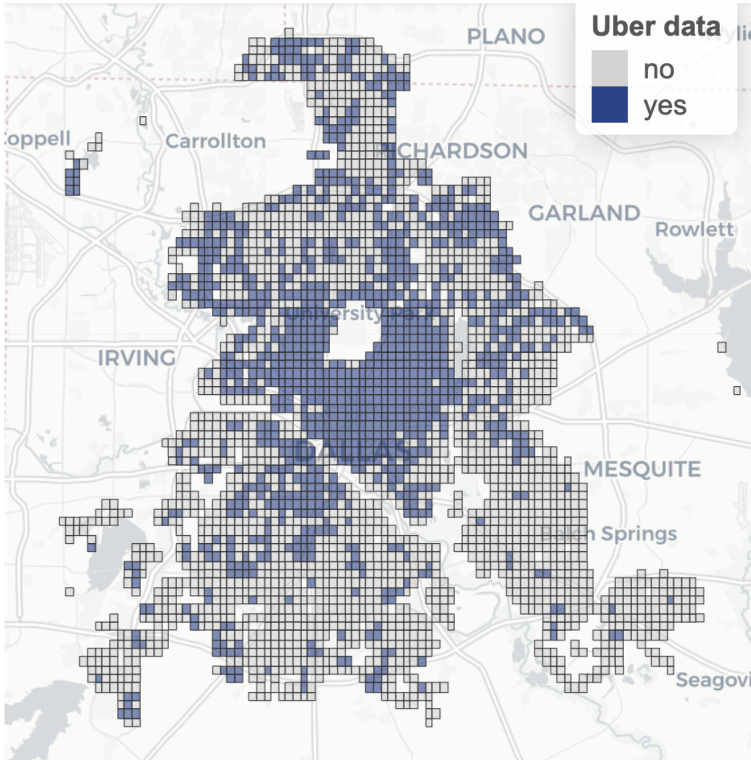
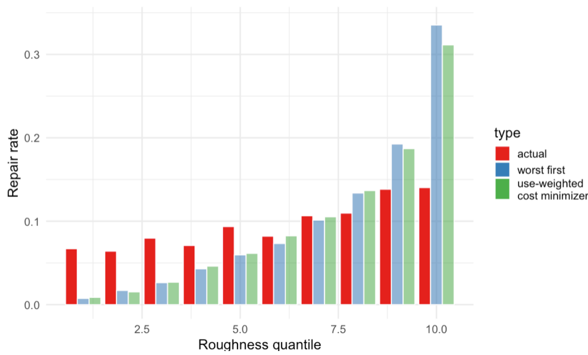


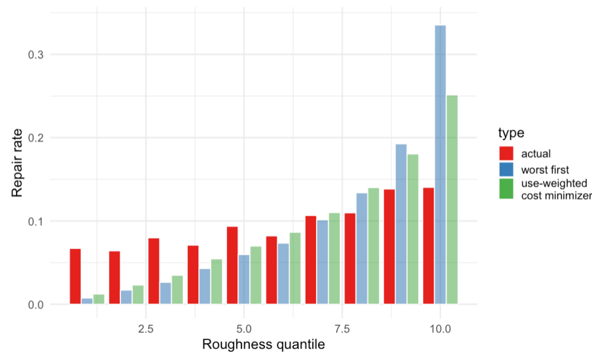
Figure A.19

Note: This figure shows the percent of girds in Dallas that have some Uber coverage, i.e. for which we have a measure of local road roughness. 37% of grids within the city boundary have some Uber coverage. For New York City, Columbus, and Portland, the numbers are 93%, 51%, and 43%.

Figure A.20: NYC Roughness Targeting Controlling for Income



(a)



(b)

Note: This figure shows the percent of roads that are repaired within each decile of roughness cost. Panel (a) repeats the graph from Figure 14, while Panel (b) shows the same results where the road usage measure (Uber traffic) has been residualized by income in the use-weighted cost minimizing counterfactual. For all graphs, arterial and highways are excluded. Income is at the grid level and imputed from ACS Census tract household medians. Roughness cost is at the grid cell level, and is the average over the segments with Uber data within the cell. Grid cells with Uber data on less than 20% of local roads are dropped.

B Appendix Tables

Table A.1: Uber Roughness Data Coverage of Road Network

	(1) With FE	(2) Total	(3) Coverage Share	(4) Coverage Share (Pop. Weighted)
Panel A. Number of Segments				
Local roads	1,268,601	27,112,992	0.047	0.14
Arterial roads	3,850,221	9,905,938	0.39	0.63
Highways	523,538	1,147,462	0.46	0.73
Panel B. Segment Length (kilometers)				
Local roads	89,274	5,607,603	0.016	0.12
Arterial roads	296,290	1,701,138	0.17	0.63
Highways	130,575	399,187	0.33	0.75

Note: This table reports raw and population-weighted coverage of the Uber road roughness data. To construct it, we start with the universe of road segments. The first three columns report the number of segments and total length with Uber road roughness fixed effects (column 1), in total (column 2), and their ratio (column 3). Column (4) restricts to observations within towns (census places) and reports the share with fixed effects using town population weights.

Table A.2: Number of Uber Trips per Segment and Road Roughness

	Predicted roughness at 20mph			
	(1)	(2)	(3)	(4)
Log trips per segment	0.0116***	0.0025***	0.0013	-0.0008
	(0.0002)	(0.0002)	(0.0008)	(0.0008)
Constant	0.2838***	0.3279***	0.3292***	0.3384***
	(0.0008)	(0.0009)	(0.0035)	(0.0035)
Uber city FE		Yes		Yes
Sample: below median trips			Yes	Yes
SD of the outcome	0.16	0.16	0.17	0.17
Observations	1,829,526	1,829,524	914,586	914,584
R2	0.00	0.04	0.00	0.04

Note: This table reports the correlation between log number of Uber trips and predicted road roughness at 20mph for all local road segments in our data. The data covers 240 Uber cities. Only road segments with at least 50 trips per segment are included. Column 3 and 4 restrict to the sample of road segments with below-median number of trips per segment (103 trips). Robust standard errors in parentheses, *p<0.1; **p<0.05; ***p<0.01.

Table A.3: Cross-Validation of Segment-Level Uber Road Roughness and Speed (Chicago)

	Predicted Roughness			Log Speed		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Outcome Using April Data						
Predicted Roughness (March data)	0.91***	0.97***	0.93***			
	(0.01)	(0.00)	(0.00)			
Log Speed (March data)				1.00***	0.99***	0.99***
				(0.00)	(0.00)	(0.00)
Road type	highway	arterial	local	highway	arterial	local
Observations	8,823	225,035	156,144	8,823	225,035	156,144
Adj R2	0.82	0.90	0.84	0.99	0.99	0.98
Panel B: Outcome Using August Data						
Predicted Roughness (March data)	0.90***	0.94***	0.90***			
	(0.01)	(0.00)	(0.00)			
Log Speed (March data)				0.99***	0.99***	0.98***
				(0.00)	(0.00)	(0.00)
Road type	highway	arterial	local	highway	arterial	local
Observations	8,816	225,742	159,632	8,816	225,742	159,632
Adj R2	0.80	0.86	0.80	0.99	0.98	0.97

Note: This table reports cross-validation results in Chicago for predicted roughness and log median speed using March and April data (Panel A) and August data (Panel B). Each column reports results from a regression at the road segment level where the outcome is estimated only using the later data (April or August 2021), and the explanatory variable is estimated only using March 2021 data. Predicted roughness is computed using 20mph for local roads, 32mph for arterial roads, and 48mph for highways. Robust standard errors in parentheses, *p<0.1; **p<0.05; ***p<0.01 .

Table A.4: Split Sample Correlation Between Predicted Roughness and Speed (Chicago)

	Median Speed								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Train	Test	Off-peak	Train	Test	Off-peak	Train	Test	Off-peak
Predicted Roughness (Training data)	-29.0***	-29.4***	-29.3***	-27.6***	-30.5***	-30.4***	-20.5***	-21.2***	-21.1***
	(1.0)	(1.1)	(1.1)	(0.2)	(0.2)	(0.2)	(0.1)	(0.2)	(0.2)
Constant	59.9***	59.3***	59.3***	40.6***	40.8***	40.7***	26.9***	26.4***	26.3***
	(0.5)	(0.5)	(0.5)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
Road type	highway	highway	highway	arterial	arterial	arterial	local	local	local
Observations	9,021	9,021	9,021	219,661	219,661	219,661	94,177	94,177	94,177
Adj R2	0.11	0.10	0.10	0.21	0.22	0.22	0.22	0.21	0.21

Note: This table reports same- and split-sample regressions of road segment median speed on predicted roughness, including using only off-peak speed measurement. This table uses data from August 2021. We first divide the data into a 75% sample used for estimation (training), and a 25% hold-out sample used for testing. We compute median road segment speed in both samples. We further compute median speed only during off-peak hours () in the testing sample. In all the regressions, we regress median speed on predicted roughness. In odd columns, we compute median speed in the same data that we use for estimating roughness. In odd columns, we use the hold-out sample to compute speed. Predicted roughness is computed using 20mph for local roads, 32mph for arterial roads, and 48mph for highways. Robust standard errors in parentheses, *p<0.1; **p<0.05; ***p<0.01 .

Table A.5: The Impact of Roughness on Driver Speed on Local Roads at Town Borders (linear)

	Log speed (mph)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log roughness slope	-0.455***	-0.406***	-0.327***	-0.327***	-0.314***	-0.314***	-0.298***	-0.298***
	(0.020)	(0.013)	(0.015)	(0.015)	(0.017)	(0.017)	(0.020)	(0.020)
Log speed limit	0.237***		0.188***		0.192**		0.044	
	(0.037)		(0.053)		(0.063)		(0.022)	
Sample:	Town	Town	Borders	Borders	Borders	Borders	Borders	Borders
Sample restriction:			$<$ 1km	$<$ 1km	$<$ 500m	$<$ 500m	$<$ 1km	$<$ 1km
Uber City FE	Yes	Yes						
Border pair FE			Yes	Yes	Yes	Yes	Yes	Yes
Distance to border controls							Yes	Yes
Uber cities	83	170	73	73	72	72	73	73
Towns	1,509	3,752	1,245	1,245	1,209	1,209	1,245	1,245
Border pairs	0	0	1,366	1,366	1,285	1,285	1,366	1,366
N	1,509	3,752	2,732	2,732	2,570	2,570	2,732	2,732

Note: Version of Table 6 with a linear specification in log roughness slope $\log(\gamma_r)$ and log speed limit $\log(s_k^{\mathrm{LIM}})$. Rather than using town or border-side dummies as instruments, here we take averages at that level and run an OLS regression. In column 1, we average log speed and log roughness slope at the town level and then run the regression. (The speed limit variable measures median speed limit at the town level so it is already constant). In column 2 and 3, we take the average at the border-side level. In column 4, we first residualize log speed and log roughness slope on distance to the border with a separate coefficient for each border-side. We then average at the border-side level. Standard errors, adjusted for clustering at the Uber city level (column 1) and border pair level (columns 2-4) in parentheses. Standard errors clustered at the level of Uber cities in parentheses, *p<0.1; **p<0.05; ***p<0.01.

Table A.6: The Impact of Roughness on Driver Speed on Local Roads at Town Borders (semi-elasticities)

	Log speed (mph)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Local roughness (z-score)	-0.155***	-0.146***	-0.094***	-0.094***	-0.103***	-0.102***	-0.226***	-0.225***
	(0.010)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.013)	(0.013)
Log speed limit	0.251***		0.202***		0.204**		0.052*	
	(0.040)		(0.052)		(0.062)		(0.023)	
Sample:	Town	Town	Borders	Borders	Borders	Borders	Borders	Borders
Sample restriction:			< 1km	< 1km	< 500m	< 500m	< 1km	< 1km
Uber City FE	Yes	Yes						
Border pair FE			Yes	Yes	Yes	Yes	Yes	Yes
Distance to border controls							Yes	Yes
Uber cities	83	170	73	73	72	72	73	73
Towns	1,509	3,752	1,245	1,245	1,209	1,209	1,245	1,245
Border pairs	0	0	1,366	1,366	1,285	1,285	1,366	1,366
N	1,509	3,752	2,732	2,732	2,570	2,570	2,732	2,732

Note: Version of Table A.5 using as dependent variable the z-score for predicted road roughness at 20 miles per hour. Standard errors clustered at the level of Uber cities in parentheses, *p<0.1; **p<0.05; ***p<0.01.

Table A.7: The Impact of Roughness on Driver Speed on Local Roads in Chicago using Repaving Events (linear IV)

	Log roughness slope		Log speed (mph)		
	(1)	(2)	(3)	(4)	(5)
Log roughness slope			-0.145***	-0.145***	-0.141***
			(0.026)	(0.026)	(0.030)
Repaved	0.206***	0.206***	-0.009	-0.009	
	(0.036)	(0.036)	(0.012)	(0.012)	
Post repaving	-0.025***	-0.025***	-0.014***	-0.014***	0.000
	(0.007)	(0.007)	(0.002)	(0.002)	(NaN)
Post repaving × Repaved	-0.191***	-0.191***			
	(0.020)	(0.020)			
Log speed limit	0.156*		0.010		
	(0.077)		(0.022)		
Repaving Event FE	Yes	Yes	Yes	Yes	Yes
Road Segment FE					Yes
Repaving Event × Post FE					Yes
Estimator	OLS	OLS	IV	IV	IV
Repaving Events	611	611	611	611	611
N	19,878	19,878	19,878	19,878	19,878

Note: Version of Table 7 with a linear 2SLS specification. The estimating equation is a version of equation (4.9), where $\log(\gamma_r)$ and $\log(s_k^{\text{LIM}})$ enter additively instead of the composite term $\log\left(\gamma_r + \frac{\theta}{s_k^{\text{LIM}}}\right)$. Column 1 and 2 report the first stage, with and without controlling for the segment speed limit. In columns 3-5, log roughness slope is instrumented using the Post repaving × Repaved interaction. Standard errors, adjusted for clustering at the road repaving event level in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A.8: The Impact of Roughness on Driver Speed on Local Roads in Chicago using Repaving Events

	Dependent variable:
	Log speed (mph)
	(1)
Log speed costs $\frac{\sigma}{\sigma+1}$	-0.15
	[-0.20, -0.08]
Inverse speed limit θ	0.01
	[0.01, 0.01]
Repaved road	-0.01
	[-0.03, 0.01]
Post repaving	-0.01
	[-0.02, -0.01]
Welfare factor $1 + \frac{1}{\sigma}$	6.66
	[4.98, 12.02]
Repair events	611
Observations	19,878

Note: Version of Table 7 using $\theta = 0.01$. In parentheses, 95% confidence intervals based on bootstrap at the repaving event level.

Table A.9: Inspection Failure Rate and Road Roughness By Indicator

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Brakes	Front	Suspensions	Frame	Muffler	Bumper	Tires	Wipers	Windshield	Mirror
Local Roughness (z-score)	0.022	0.084	0.072	0.009	0.076	0.028	0.084	0.052	0.018	0.022
	(0.055)	(0.088)	(0.077)	(0.009)	(0.043)	(0.036)	(0.084)	(0.047)	(0.011)	(0.015)
Mean Outcome	0.18	0.97	0.70	0.03	0.45	0.21	1.10	0.47	0.11	0.06
Observations	926	926	926	926	926	926	926	926	926	926

Note: This table reports the correlation between inspection failure rates and local road roughness in the same town. The data is at the inspection station level. The outcome is the level of failures per inspection, for the respective category, namely brakes, front end, steering and suspension frame, muffler and exhaust system, bumpers/fenders/exterior sheet metal, tires, windshield wipers and cleaner, windshield, and rear view mirror. The last three indicators are the placebo indicators. Robust standard errors in parentheses, *p<0.1; **p<0.05; ***p<0.01.

Table A.10: Tract Coverage Robustness Check

	Dependent variable:					
	Cost (USD per mile)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: > 5% coverage						
ln median income	-0.037^{***}	-0.048^{***}	-0.014^{***}			
	(0.001)	(0.001)	(0.002)			
fraction Black				0.075^{***}	0.080^{***}	0.009^{***}
				(0.003)	(0.003)	(0.003)
Observations	23,826	23,965	23,965	23,939	24,080	24,080
Panel B: > 10% coverage						
ln median income	-0.029^{***}	-0.041^{***}	-0.009^{***}			
	(0.002)	(0.002)	(0.002)			
fraction Black				0.059^{***}	0.065^{***}	0.004
				(0.003)	(0.003)	(0.003)
Observations	18,535	18,649	18,649	18,636	18,752	18,752
Panel C: > 15% coverage						
ln median income	-0.021^{***}	-0.032^{***}	-0.006^{***}			
	(0.002)	(0.002)	(0.002)			
fraction Black				0.048^{***}	0.052^{***}	0.003
				(0.004)	(0.004)	(0.004)
Observations	15,056	15,152	15,152	15,148	15,246	15,246
Climate controls	Yes			Yes		
MSA Fixed effects		Yes			Yes	
Town Fixed effects			Yes			Yes

Note: This table reports results from regressions following the specifications in Table 9 for income and the faction of households that are Black. Here, we adjust the data sample by removing any tract for which the percent of road segments we have roughness for is below a threshold. In the first panel, we remove tracts with less than 5% coverage of segment roughness; in the second, less than 10; in the final panel, less than 15%. The coefficients persist as we restrict the sample, but the magnitude shrinks noticeably. Given the evidence in Figure A.5 that there is no relationship between roughness and number of Uber observations, we interpret these results as suggestive that the relationship between roughness and income/race is strongest in tracts with the worst Uber coverage. Robust standard errors in parentheses, *p<0.1; **p<0.05; ***p<0.01.

Table A.11: Road Expenditure Correlates

	Dependent variable:					
	ln expenditure per capita		ln expenditure			
	(1)	(2)	(3)	(4)	(5)	(6)
ln median income	0.528***		0.045		0.068	
	(0.105)		(0.148)		(0.119)	
fraction Black		−0.341*		0.184		−0.098
		(0.191)		(0.274)		(0.189)
fraction Hispanic		−1.590***		−0.701**		−0.891***
		(0.205)		(0.294)		(0.239)
fraction Asian		−1.783***		−0.615		−2.074***
		(0.403)		(0.577)		(0.384)
ln miles of local road					0.205	0.013
					(0.171)	(0.166)
ln population					0.362***	0.508***
					(0.116)	(0.111)
miles to CBD					0.025	−0.009
					(0.056)	(0.057)
ln employment					0.273***	0.310***
					(0.047)	(0.048)
ln area (miles^2)					0.404***	0.414***
					(0.111)	(0.110)
fraction drive to work					−2.187***	−1.550***
					(0.417)	(0.434)
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,223	1,223	1,223	1,223	1,223	1,223
Adjusted R^2	0.215	0.239	0.157	0.161	0.638	0.648

Note: This table reports results from regressions of road expenditure on town level covariates. The first two columns have log expenditure per capita as the outcome. The last four columns have log expenditure as the outcome. Note that since we include log population as a covariate in the last two columns, using the per capita measure as the dependent variable would produce the same coefficients on all the covariates except log population. Expenditure per capita is each local government's reported direct expenditures on "construction and maintenance of roads, sidewalks and bridges; street lighting; snow removal; highway engineering, control, and safety." County and township spending are distributed to towns based on shared land area. Robust standard errors in parentheses, *p<0.1; **p<0.05; ***p<0.01.

Table A.12: Town Coverage Robustness Check

	Dependent variable:						
	Cost (USD per mile)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: no threshold							
ln median income	-0.045^{***}	-0.045^{***}	-0.027^{***}				-0.033^{***}
	(0.004)	(0.004)	(0.005)				(0.005)
fraction Black				0.086^{***}	0.086^{***}	0.056^{***}	0.054^{***}
				(0.010)	(0.010)	(0.009)	(0.011)
ln expenditure per capita		-0.0003	0.001		-0.002	0.001	
		(0.001)	(0.001)		(0.001)	(0.001)	
MSA Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls			Yes			Yes	
Observations	3,142	3,142	3,142	3,142	3,142	3,142	3,142
Panel B: 5% coverage							
ln median income	-0.054^{***}	-0.056^{***}	-0.036^{***}				-0.044^{***}
	(0.005)	(0.005)	(0.005)				(0.005)
fraction Black				0.080^{***}	0.081^{***}	0.062^{***}	0.044^{***}
				(0.009)	(0.009)	(0.008)	(0.010)
ln expenditure per capita		0.004^{***}	0.005^{***}		0.001	0.004^{***}	
		(0.001)	(0.001)		(0.001)	(0.001)	
MSA Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls			Yes			Yes	
Observations	1,902	1,902	1,902	1,902	1,902	1,902	1,902
Panel C: 15% coverage							
ln median income	-0.044^{***}	-0.047^{***}	-0.025^{***}				-0.033^{***}
	(0.007)	(0.007)	(0.007)				(0.008)
fraction Black				0.065^{***}	0.065^{***}	0.060^{***}	0.041^{***}
				(0.012)	(0.012)	(0.010)	(0.013)
ln expenditure per capita		0.005^{**}	0.005^{***}		0.003	0.005^{**}	
		(0.002)	(0.002)		(0.002)	(0.002)	
MSA Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls			Yes			Yes	
Observations	866	866	866	866	866	866	866

Note: This table reports results from regressions following the specifications in Table 12, for income and the faction of households that are Black. Here, we adjust the data sample by removing any town for which the percent of road segments we have roughness for is below a threshold. In the first panel, we remove no towns; in the second, less than 5; in the final panel, less than 15% (in Table 12, we remove towns with less than 10% coverage). The coefficients persist as we restrict the sample, but the magnitude shrinks noticeably. Given the evidence in Figure A.5 that there is no relationship between roughness and number of Uber observations, we interpret these results as suggestive that the relationship between roughness and income/race is strongest in tracts with the worst Uber coverage. Robust standard errors in parentheses, *p<0.1; **p<0.05; ***p<0.01.

Table A.13: NYC Repair Targeting with Sub-Sample

	Dependent variable:	
	Repair rate (percent road-miles repaved) NYC	
	(1)	(2)
local roughness (z score)	0.015***	0.015***
	(0.005)	(0.005)
ln road miles	0.004	0.002
	(0.004)	(0.005)
ln population	0.007**	0.011***
	(0.003)	(0.004)
ln miles to CBD	−0.040***	−0.048***
	(0.006)	(0.008)
fraction Black	−0.031***	−0.028***
	(0.011)	(0.011)
ln median income	0.005*	0.002
	(0.003)	(0.003)
ln Uber volume		−0.005
		(0.006)
fraction drive to work		0.025
		(0.022)
Constant	0.080**	0.121**
	(0.038)	(0.052)
Observations	2,708	2,708
Adjusted R^2	0.043	0.043

Note: This table reports results from regressions following the specifications in Table 13 for New York City, but where the data has been subsampled to match the distribution of Uber coverage across grids in Dallas. Specifically, for Dallas and New York City, we group grids into 5 coverage quantiles (conditional on > 0 Uber data) and get the average coverage within each quantile. For each NYC grid g we get the hypothetical length of road it would need to have Uber coverage on to match the mean Dallas coverage in the same quantile. We define $p_g = \min(\text{the hypothetical length/the true Uber segment length}, 1)$ Then we keep or discard each Uber segment in g with probability p_g.

Table A.14: Repair Targeting with Varying Coverage Thresholds

	Dependent variable:							
	Repair in grid							
	NYC		Dallas		Columbus		Portland	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: > 20% coverage								
local roughness (z score)	0.026***	0.026***	−0.008	−0.014	0.007	0.004	−0.002	−0.002
	(0.006)	(0.006)	(0.011)	(0.011)	(0.008)	(0.008)	(0.002)	(0.002)
ln road miles	0.004	0.004	0.035***	0.030***	−0.024***	−0.023***	0.002	0.002
	(0.004)	(0.004)	(0.008)	(0.008)	(0.005)	(0.005)	(0.002)	(0.002)
ln population	0.005	0.008**	−0.006	0.002	0.016**	0.013**	0.0003	0.0003
	(0.003)	(0.004)	(0.007)	(0.007)	(0.007)	(0.007)	(0.001)	(0.002)
ln miles to CBD	−0.035***	−0.040***	−0.0004	−0.035**	−0.020**	0.001	0.001	0.004
	(0.006)	(0.008)	(0.012)	(0.015)	(0.008)	(0.009)	(0.002)	(0.005)
fraction Black	−0.035***	−0.032***	−0.033	−0.029	0.001	0.025	−0.008	−0.007
	(0.011)	(0.011)	(0.040)	(0.041)	(0.022)	(0.023)	(0.035)	(0.035)
ln median income	0.004	0.002	0.009	0.010	0.022*	0.052***	−0.0002	0.001
	(0.003)	(0.003)	(0.007)	(0.007)	(0.012)	(0.014)	(0.004)	(0.004)
ln Uber volume		−0.009		−0.050***		0.011		−0.002
		(0.007)		(0.014)		(0.009)		(0.004)
fraction drive to work		0.009		0.336**		−0.238***		−0.016
		(0.023)		(0.131)		(0.060)		(0.018)
Constant	0.087**	0.143**	0.068	0.025	−0.270*	−0.477***	0.003	−0.002
	(0.038)	(0.056)	(0.074)	(0.167)	(0.150)	(0.158)	(0.045)	(0.054)
Observations	2,732	2,732	568	568	618	618	233	233
Panel B: > 30% coverage								
local roughness (z score)	0.029***	0.029***	−0.013	−0.021	0.013	0.011	−0.003	−0.003
	(0.006)	(0.007)	(0.014)	(0.014)	(0.012)	(0.012)	(0.002)	(0.002)
ln road miles	0.003	0.002	0.038***	0.036***	−0.031***	−0.032***	0.002	0.002
	(0.004)	(0.004)	(0.009)	(0.009)	(0.006)	(0.006)	(0.001)	(0.001)
ln population	0.006*	0.009**	−0.011	0.0004	0.016*	0.011	0.001	0.001
	(0.004)	(0.004)	(0.008)	(0.009)	(0.009)	(0.008)	(0.001)	(0.001)
ln miles to CBD	−0.036***	−0.042***	0.001	−0.038**	−0.025**	−0.001	−0.0003	−0.0001
	(0.006)	(0.008)	(0.014)	(0.018)	(0.011)	(0.012)	(0.002)	(0.004)
fraction Black	−0.032***	−0.029***	−0.065	−0.066	−0.021	0.014	−0.018	−0.018
	(0.011)	(0.011)	(0.049)	(0.051)	(0.030)	(0.031)	(0.032)	(0.032)
ln median income	0.004	0.002	0.013*	0.011	0.013	0.053***	0.003	0.003
	(0.003)	(0.003)	(0.007)	(0.008)	(0.015)	(0.019)	(0.003)	(0.004)
ln Uber volume		−0.011		−0.065***		0.022*		0.0004
		(0.008)		(0.017)		(0.012)		(0.003)
fraction drive to work		0.012		0.365**		−0.270***		−0.001
		(0.023)		(0.155)		(0.077)		(0.017)
Constant	0.083**	0.152**	0.066	0.089	−0.162	−0.505**	−0.032	−0.035
	(0.040)	(0.060)	(0.078)	(0.192)	(0.188)	(0.202)	(0.039)	(0.048)
Observations	2,596	2,596	427	427	396	396	160	160
Note:								*p<0.1; **p<0.05; ***p<0.01

Note: This table reports results from regressions following the specifications in Table 13. Here, we vary the Uber coverage threshold under which we remove any grid from the sample. In the first panel, we remove grids with less than 20% coverage of segment roughness; in the second, we remove grids with less than 30% coverage of segment roughness. There is a trade-off to restricting the sample. Requiring higher coverage reduces measurement error, but also changes the sample if there is selection in where Uber riders travel.

Table A.15: Repair Targeting vs. Counterfactual Targeting with Varying Coverage Thresholds

	Dependent variable:							
	Repair rate (percent road-miles repaved)							
	NYC		Dallas		Columbus		Portland	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: > 0 coverage								
worst first	0.090***		0.017		−0.019		0.002	
	(0.020)		(0.028)		(0.032)		(0.014)	
use-weighted		0.162***		−0.073		−0.005		−0.008
		(0.022)		(0.063)		(0.048)		(0.024)
ln road miles	0.012***	0.010***	0.037***	0.037***	−0.015***	−0.015***	0.001	0.001
	(0.003)	(0.003)	(0.006)	(0.006)	(0.003)	(0.003)	(0.001)	(0.001)
Constant	0.074***	0.070***	0.125***	0.127***	0.021***	0.021***	0.005***	0.005***
	(0.004)	(0.004)	(0.005)	(0.005)	(0.003)	(0.003)	(0.001)	(0.001)
Observations	2,974	2,974	1,086	1,086	1,249	1,249	619	619
Adjusted R^2	0.010	0.020	0.039	0.039	0.019	0.019	−0.002	−0.002
Note:								*p<0.1; **p<0.05; ***p<0.01
Panel B: > 30% coverage								
worst first	0.099***		−0.018		0.018		−0.011	
	(0.022)		(0.059)		(0.114)		(0.033)	
use-weighted		0.164***		−0.155		0.094		−0.032
		(0.024)		(0.119)		(0.142)		(0.067)
ln road miles	0.012***	0.010***	0.040***	0.042***	−0.023***	−0.023***	0.002*	0.002*
	(0.003)	(0.003)	(0.008)	(0.008)	(0.006)	(0.006)	(0.001)	(0.001)
Constant	0.075***	0.071***	0.126***	0.130***	0.017**	0.016**	0.003**	0.003**
	(0.004)	(0.004)	(0.010)	(0.010)	(0.007)	(0.007)	(0.001)	(0.001)
Observations	2,650	2,650	427	427	396	396	160	160
Adjusted R^2	0.011	0.021	0.056	0.059	0.033	0.034	0.011	0.012
Note:								*p<0.1; **p<0.05; ***p<0.01

Note: This table reports results from regressions following the specifications in Table 14, but where we adjust the data sample by removing any gird for which the percent of road segments we have roughness for is below different. In Table 14, the threshold is 20%. In the first panel, we remove no grids; in the second, we remove grids with less than 30% coverage of segment roughness. There is a trade-off to restricting the sample. Requiring higher coverage reduces measurement error, but also changes the sample if there is selection in where Uber riders travel.

C Data Appendix

C.1 Geographic data.

We obtain boundaries for census tracts, counties, towns (Census places), CBSAs, and states, as well as geographic crosswalks between them, from the US Census' TIGER/Line Shapefiles.

C.2 Census data, 2015-2019.

We use data on population, income, race, and ethnicity from the 2019 American Community Survey 5-Year estimates (2015-2019) at the Census tract and Census place level. We assign tracts to towns (Census Designated Places) using crosswalks from the Census. We join this with data from the 2018 Zip Codes Business Patterns. Zip Code Tabulation Areas are matched to towns by share of overlapping area.

C.3 Climate data for the entire United States, 2016-2022.

We use county-level monthly maximum temperatures, minimum temperatures, and precipitation rates from the National Oceanic and Atmospheric Administration (NOAA). For each county, we take a five-year average of each outcome.

C.4 IRI for the entire United States, 2017-2018 (highways).

We use International Roughness Index (IRI) data provided by the federal Department of Transport's Highway Monitoring Performance System (HPMS). The latest available data as of 2023 is from 2017-2018. We download the data from https://www.fhwa.dot.gov/policyinformation/hpms/shapefiles.cfm. The data includes a shapefile of road segments and IRI measurement. We include all roads in the National Highway System for IRI data, while for our Uber data, we only those roads that our data classifies as "highways".

C.5 International Roughness Index (IRI) in Cook county, 2018 (arterial roads, highways).

We use data on segment-level road roughness measured as the International Roughness Index (IRI) for Cook County provided by the Illinois Department of Transportation (IDOT) and collected as part of the Transportation Asset Management Plan (TAMP). Road quality information is collected 1) to prioritize highway rehabilitation needs, and 2) for incorporation into the Highway Performance Monitoring System (HPMS), the Federal Highway Administration's (FHWA) database used to report on the national road network and apportion federal funding to states for transportation needs. The IRI, along with other road measures, is collected by vans

with laser sensors and on-board computers. This data is primarily available for highways and arterial roads, with very sparse coverage of residential roads.
Data on the road network is published with respect to a basemap, which provides the coordinates for each road segment on a model of the Earth. Different basemaps differ in the precise location of roads. Therefore, comparing across datas ets requires matching segments. We match OSM basemap road segments in Chicago to the TAMP basemap road segments based on the distance separating them, a fuzzy name match of the road names, and the angle generated by the intersection of tangent lines at the nearest point between the segments.
For each OSM segment, we assign it an IRI roughness level equal to the weighted average of overlapping TAMP segments, where the weights are the length of the overlap between segments divided by the length of the OSM segment. We also calculate a match quality measure that depends on how much TAMP segments extend away from the OSM segment, as in these cases the IRI measure for a given OSM segment may be particularly noisy.

C.6 Pavement Condition Index (PCI) in New York City, 2021.

We use data on the PCI on New York City streets provided by NYC OpenData. We limit our analysis to PCI estimates collected between June and October 2021.

C.7 Railway crossings at grade in Chicago.

We obtained the universe of at-grade railway crossings in Illinois from 2017 from the Illinois Department of Transportation, from https://gis-idot.opendata.arcgis.com/ (last accessed June 26, 2023).

C.8 Road resurfacing in Chicago, 2021.

We collected data on road resurfacing in Chicago. This data is based on moratoriums on street work. The City of Chicago imposes an increased permit fee on anyone who wishes to excavate a given street within five years of its last resurfacing. The city consequently provides the start and end address of the moratorium section of road and the moratorium start and end date. We infer the approximate resurfacing completion date by subtracting five years from the expiration date.[45] As we discuss when we return to resurfacing, these inferences are likely to contain some spatial and temporal noise.

[45]There are some concerns with the validity of the provided dates. Although the public website says the moratorium length is 7 years for resurfacing, discussions with the Chicago Department of Transportation confirmed that the correct length is 5 years. This is, indeed, the modal length between start and end dates in the data. However, there are also a number of odd lengths, possibly due to the start date corresponding to the data entry date rather than the construction date. For this reason, we follow the advice of the Chicago DOT and infer completion date as 5 years prior to the moratorium expiration date.

For this analysis, we compile the set of road resurfacing moratoriums completed in Chicago from May through mid-July, 2021. We match the repaired sections of road to our data by using the Google Geocoding API to turn addresses to coordinates, and an algorithm based on distance and a fuzzy street name match. These road segments are part of the treated or repaved group.

C.9 Road resurfacing in New York, Dallas, Portland, and Columbus, 2021-2022.

We collect data on resurfacing in New York City, Dallas, Portland, and Columbus from an argGIS map published on each respective city's public website. We use data from September 2021 through December 2022 for New York, Dallas, and Portland, and data from only 2022 for Columbus, which does not publish more precise data on times. To remove highway and arterial repairs, we drop any repaired segments that are contained by a 25-meter buffer of OSM basemap arterial and highway roads.

We interpolate Census variables defined at the Census tract level to the grid cell level as follows. First, we assume that variables are uniformly distributed within each census tract, so we can use the area of the intersection between a tract and a grid cell. For "mean" variables, such as income, we take the weighted mean of the variable for the census tracts that intersect a given grid cell, with weights given by the area of the intersection. For aggregate variables, such as total population, we first apportion a census tract's value to all of its intersections with grid cells, proportional to the areas of the intersections. Next, for each grid cell, we take the sum of the variable for the intersections.

C.10 Speed limit data, 2023.

We collect road-segment level data on official posted speed limits for a random sample of road segments used in our town border analysis and for all road segments in our Chicago road resurfacing analysis. We use the HERE.com API. (HERE.com is a mapping data platform.) For the town border analysis, we compute the median speed limit at the town level.

C.11 Vehicle inspection failure rates data for Massachusetts, 2021-2022.

We obtained data on non-commercial vehicle inspection failure rates in Massachusetts from the Massachusetts Department of Transport (MassDOT/Registry of Motor Vehicle). We obtain data at the level of each inspection location (station) with the total number of inspections, average vehicle age, and total number of failures by category for 10 categories, for the period May 2021 to April 2022. We pre-selected seven "main" forms of inspection failure that we hypothesized could be linked to vehicle damage due to rough roads: brakes, front end, steering and suspension frame, muffler and exhaust system, bumpers/fenders/exterior sheet metal, and tires. We also pre-selected three "placebo" types of failure that we thought unlikely to be related to rough roads: windshield wipers and cleaner, windshield, and rearview mirror.

C.12 Vehicle Crash Fatalities for the entire United States, 2021.

We use data on the universe of vehicle crash fatalities from the Fatality Analysis Reporting System (FARS) published by the National Highway Traffic Safety Administration (NHTSA). We accessed data for 2021 from https://www.nhtsa.gov/file-downloads?p=nhtsa/downloads/FARS/.

C.13 Policymaker survey of road repair data and decisions.

The goal of this survey is to obtain data on road repaving strategies from municipalities across the US. The list of cities is derived from the 2010 Census and excludes Census Designated Places, as well as 15 towns for miscellaneous reasons (no contact information online, no department of public works, county responsible for road repaving, etc.). The sample includes a national sample, with a 16.3% response rate and a Massachusetts sample for which more follow-ups were conducted, achieving a 63% response rate.

The contact protocol includes randomizing email send time, spacing 3 business days between emails and 4 days between emails and calls, and ensuring that every business day, 25 new cities and 25 previously contacted cities are emailed the survey.

We report here results corresponding to Tables 15 and 16 for cities that responded later to our survey (Table A.16).

Table A.16: Survey Responses by Late Response

	Dependent variable:					
	In-house Survey	Private Survey	Officials	Engineering	Residents	Share Surveyed
	(1)	(2)	(3)	(4)	(5)	(6)
Responds Late	−0.38	0.20	0.00	−0.39*	−0.29	−18.83
	(0.22)	(0.14)	(0.00)	(0.21)	(0.20)	(17.20)
Constant	0.60***	0.80***	0.00	0.50***	0.40**	70.50***
	(0.15)	(0.10)	(0.00)	(0.14)	(0.14)	(11.84)
Observations	19	19	19	19	19	19

	Dependent variable:						
	Traffic	Formula Road Conditions	Formula Other	Utility Work	Elected	Citizen Complaints	Accessibility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Responds Late	−0.06	0.16	−0.08	0.26	0.33**	0.00	−0.10
	(0.24)	(0.24)	(0.21)	(0.23)	(0.16)	(0.00)	(0.11)
Constant	0.50***	0.40**	0.30*	0.30*	−0.00	0.00	0.10
	(0.17)	(0.16)	(0.15)	(0.16)	(0.11)	(0.00)	(0.07)
Observations	19	19	19	19	19	19	19

	Dependent variable:	
	Gets Required Resurfacing	Resurfacing Spending Share
	(1)	(2)
Responds Late	−22.50*	8.44
	(12.42)	(12.31)
Constant	47.50***	66.00***
	(8.54)	(8.48)
Observations	19	19

Note: These tables analyze heterogeneity of results from Tables 15 and 16. We sent reminder emails and calls to towns that did not fill out the survey. The variable “Responds Late” is a dummy for above median response time.

D Model Appendix

D.1 The Elasticity σ and the Marginal Cost of Roughness

In this section, we derive the impact of a higher elasticity σ on the marginal effect of road roughness on total costs and show that this effect is ex-ante ambiguous. We perform this comparative static holding fixed the optimal speed at the median value of road roughness γ_r. This normalization is motivated by the two empirical settings we consider, the nation-wide town border sample and the Chicago road resurfacing sample. In both cases, the median speed is approximately 20 mph. Meanwhile, we find a higher speed elasticity in the case of town borders than for Chicago resurfacing events.

We re-write the speed cost shifter as $\beta_{ri} = \widetilde{\beta}_{ri}\overline{\beta}$, where $\overline{\beta}$ is notation for the median speed cost shifter for a segment with median roughness. Using equation (4.3) for optimal speed, we use the normalization

$$(\overline{s})^{-(1+\sigma)} = \frac{\overline{\beta}}{v_i}\Big(\overline{\gamma} + \frac{\theta}{\overline{s}^{\text{LIM}}}\Big)^{\sigma},$$

where $\overline{\gamma}$ is median roughness, $\overline{s}^{\text{LIM}}$ is median speed limit (conditional on roughness), and $\overline{s}$ is optimal speed for a road segment with the given roughness, speed cost shock, and speed limit. Re-arranging, we get the following expression for a road segment's cost shock after we take into account the normalization.

$$\beta_{ri} = \widetilde{\beta}_{ri}\frac{v_i}{\overline{s}}\Big(\overline{\gamma s} + \theta\frac{\overline{s}}{\overline{s}^{\text{LIM}}}\Big)^{-\sigma}.$$

Substituting β_{ri} back into the definition of costs (4.1), we obtain

$$c_{ri}(s) = \frac{v_i}{s} + \frac{v_i}{\overline{s}}\frac{\widetilde{\beta}_{ri}}{\sigma}\left(\frac{\gamma_r s + \theta\frac{s}{s_r^{\text{LIM}}}}{\overline{\gamma s} + \theta\frac{\overline{s}}{\overline{s}^{\text{LIM}}}}\right)^{\sigma}. \tag{D.1}$$

We are interested in how σ affects the marginal impact of higher road roughness. This is given by the cross derivative $\frac{d}{d\sigma}\frac{dc^*_{ri}}{d\gamma_r}$, where c^*_{ri} is cost at the optimal speed. Using the envelope theorem, the inner derivative is the same as differentiating the cost c_{ri} with respect to γ_r, holding the optimal speed s^*_{ri} fixed. We obtain

$$\frac{dc^*_{ri}}{d\gamma_r} = \frac{dc_{ri}}{d\gamma_r} = \frac{v_i}{\overline{s}}\frac{\widetilde{\beta}_{ri}}{\sigma}\left(\frac{\gamma_r s + \theta\frac{s}{s_r^{\text{LIM}}}}{\overline{\gamma s} + \theta\frac{\overline{s}}{\overline{s}^{\text{LIM}}}}\right)^{\sigma}\left(\gamma_r + \frac{\theta}{s_r^{\text{LIM}}}\right)^{-1}.$$

The derivative of this expression with respect to σ is positive if and only if the term in brackets is above 1, namely:

$$\frac{d}{d\sigma}\frac{dc^*_{ri}}{d\gamma_r} > 0 \Leftrightarrow \gamma_r s + \theta\frac{s}{s_r^{\text{LIM}}} > \overline{\gamma s} + \theta\frac{\overline{s}}{\overline{s}^{\text{LIM}}}.$$

This condition is ex-ante ambiguous and will depend on the specific model and road segment parameters.